



저작자표시-비영리-동일조건변경허락 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.
- 이차적 저작물을 작성할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



동일조건변경허락. 귀하가 이 저작물을 개작, 변형 또는 가공했을 경우에는, 이 저작물과 동일한 이용허락조건하에서만 배포할 수 있습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

A Dissertation

For the Degree of Doctor of Philosophy

**Epigenetics of X-chromosome inactivation and
its relation to pluripotent genes in porcine early
embryogenesis**

돼지 초기 배아 발생 중 X 염색체 불활성화 현상 및
만능성 유전자와의 연관성에 관한 후생 유전학적 연구

February, 2015

By

Jae Yeon Hwang

Department of Agricultural Biotechnology

Graduate School

Seoul National University

농 학 박 사 학 위 논 문

**Epigenetics of X-chromosome inactivation and
its relation to pluripotent genes in porcine early
embryogenesis**

돼지 초기 배아 발생 중 X 염색체 불활성화 현상 및
만능성 유전자와의 연관성에 관한 후생 유전학적 연구

2015년 2월

서울대학교 대학원

농생명공학부

황 재 연

Department of Agricultural Biotechnology
Graduate School, Seoul National University

**Epigenetics of X-chromosome inactivation and its relation to
pluripotent genes in porcine early embryogenesis**

돼지 초기 배아 발생 중 X 염색체 불활성화 현상 및 만능성
유전자와의 연관성에 관한 후생 유전학적 연구

지도교수 이 창 규

이 논문을 농학박사 학위논문으로 제출함
2014 년 12 월

서울대학교 대학원
농생명공학부 동물생명공학전공
황 재 연

황재연의 농학박사 학위논문을 인준함
2014 년 12 월

위 원 장	윤 철 희	(인)
부위원장	이 창 규	(인)
위 원	김 희 발	(인)
위 원	가 학 현	(인)
위 원	김 은 배	(인)

Abstract

Epigenetics of X-chromosome inactivation and its relation to pluripotent genes in porcine early embryogenesis

Jae Yeon Hwang

Department of Agricultural Biotechnology

Graduate School

Seoul National University

X-chromosome inactivation (XCI) is an essential epigenetic process observed in female eutherian embryos to equalize expression levels of X-linked genes between male and female adults. A number of studies have been confirmed the dynamic processes and molecular mechanisms of XCI in mice. Its significance on epigenetic regulation during embryo development and as an indicator for evaluating pluripotent status in stem cells originated from pre- or post-implantation embryos has been highlighted recently. Despite its importance and broad applications in stem cells and developmental studies, the insight of epigenetic phenomenon accumulated for about fifty years has been restricted in mice. However, several studies have shown evolutionary diversities on XCI and its regulators, and suggested that analyzing XCI should be carried out with species-specific manner. Therefore, in this study, XCI in pigs were studied by 1) identifying main regulators

for XCI, *XIST* and X-chromosome inactivation center (XIC), 2) confirming epigenetic changes of XCI in preimplantation embryos, and 3) examining the relation between XCI and pluripotent genes.

The aim of the first study was the identification of *XIST* orthologs, which are the main inducer for XCI by coating inactive X-chromosome, in pigs. Although this non-coding RNA (ncRNA) and its functions for inactivating X-chromosome have been confirmed in mice and humans about twenty years ago, the ortholog in pigs has yet to be identified in pigs recently. For this, sequence comparison and expression analysis were performed in porcine embryonic fibroblasts (PEFs). The identified ncRNA was 25 Kb lengths, and its exon composition was similar to *XIST* of human and mouse. However, general sequence homologies between the transcript and *XIST* in other species were very low. Newly identified gene in the first study was expressed only in female PEFs. The gene contains four repeat regions and their monomers in the first and second repeat region were same to repeat A and repeat B in mouse *Xist*, respectively but monomers in other two repeat regions were porcine specific. Furthermore, CpG sites on its promoter were hypermethylated in male PEFs, which only have active X-chromosome. These results demonstrated that identified gene is porcine *XIST*.

X-linked genes located on close distance with *Xist* are reported to regulate XCI

in mouse. The genomic region has been called XIC. However, even though importance of the region has been highlighted because of various regulators in the region for XCI by activating (*Jpx*, *Rlim*, and *Ftx*) or repressing (*Tsix*) *Xist* expressions, the genomic region was not confirmed in pigs. Therefore, in the second study, porcine XIC was searched and expression patterns of XIC-linked genes were analyzed in porcine preimplantation embryos to examine the XCI process in pigs. XICs are generally composed to protein coding genes and ncRNAs within about 0.5 Mb and 1.1 Mb of genomic range in mice and human, respectively. The ncRNAs have been evolved by disruption of ancestral protein coding genes and insertion of mobile element during divergence of eutherian and marsupials. Proportion and transcribing strand of XIC-linked genes are conserved but the sequences of ncRNAs in XIC are less conserved among the eutherians. 1.1 Mb lengths of porcine genomic region which was a synteny with human XIC was determined to be a porcine XIC. Porcine XIC contained porcine *XIST*, and sequences of the ncRNAs in the region were less conserved compared to those of protein coding genes. These results suggest that porcine XIC shares the evolutionary properties observed in eutherian XIC. Among the XIC-linked genes, *CHIC1*, *XIST*, uncharacterized ncRNAs, *LOC102165544*, and *RLIM* were stably expressed in embryonic stages. Expression levels of protein coding genes, *CHIC1* and *RLIM* were decreased during morula to blastocyst development in females and expression levels of these genes were not different between male and female blastocysts. This would mean decrease of expression levels of the genes in female blastocysts resulted in compensating dosages of the genes between males and

females. However, other two ncRNAs, *XIST* and *LOC102165544*, showed increase of expression levels after morula stage in females and sexual dimorphic expressions were observed in blastocysts. One of the *CHIC1* promoter alleles was methylated in female blastocyst. These results mean dosage compensation of several genes would be achieved in porcine blastocysts but chromosome wide inactivation of X-chromosome would not be accomplished at the embryonic stage because both allele of *XIST* promoter were demethylated in female blastocysts.

After examining XCI in embryonic stages in pigs, relations between *OCT4*, which is a gate-keeper for maintaining pluripotency and one of the XCI regulator, and XIC-linked genes, were examined. Before examining the function of *OCT4* in XIC-linked genes, it was required to confirm the presence of *OCT4* variants in pigs because they have been reported to induce false-positive detection. Therefore, the presence of *OCT4* variants and differential expressions patterns were examined in the third study. Porcine *OCT4* transcribed additional two polyadenylated RNAs, *OCT4B* and *OCT4B1* like human *OCT4*, and they share the majority of exons consisting of *OCT4A*, which was confirmed to be a pluripotent factor, except its first exon. And *OCT4B* expressed all somatic tissues examined in the study and mainly located in cytoplasm of embryos contrary to *OCT4A* which was detected in nucleus specifically.

Next, the relation between *OCT4A* and XIC-linked genes was examined by overexpressing human *OCT4A* in developing embryos in fourth study. Contrary to the previous studies in mice, overexpression of *OCT4A* ortholog induced elevation of *XIST* expression in blastocyst when the transgene was passed to cleaved embryos. And correlative expression analysis in last study suggested that other pluripotent factors, typically *NANOG* and *REX1*, might have relation with XCI inducer, *XIST*, and its *trans* activator *RLIM*, respectively. Although more studies are needed, it was considered that molecular networking of pluripotent factors and XCI would be related during embryonic stages in pigs.

In conclusion, the present study demonstrated XCI in porcine preimplantation embryos with identifying key regulators of the process and their genomic cluster. Results in here suggested that the factors would be associated with XCI in pigs. Collectively, this study suggests new insights of porcine XCI. And also, the results obtained in the present study will advance the studies in the area of stem cells and developmental biology in pigs.

Key words: X-chromosome inactivation, X-chromosome inactivation center, *XIST*, *OCT4*, Preimplantation embryos, Blastocyst, Gene expression, Pig

Student number: 2009 - 21275

CONTENTS

ABSTRACT	i
CONTENTS	vi
LIST OF TABLES	xi
LIST OF FIGURES	xiii
LIST OF ABBREVIATIONS	xvii

CHAPTER I

<i>General Introduction</i>	1
--	----------

CHAPTER II

<i>Literature Review</i>	9
1. X-chromosome inactivation in mammals	10
1.1. X-chromosome inactivation in monotremes and marsupials	11
1.2. X-chromosome inactivation in mouse	13

1.3. X-chromosome inactivation in non-mouse eutherians	20
2. Regulators in X-chromosome inactivation	26
2.1. X-chromosome inactivation center	27
2.2 Pluripotent factor	38
3. Applications of XCI in embryo development and pluripotent stem cells	44
3.1. Aberrant incidence of XCI in developmental competence	44
3.2. Indicator for the status of pluripotency	50
4. Techniques for XCI analysis	52
4.1. Epigenetic changes of X-chromosome in its inactivation	53
4.2. Cytological analysis for evaluating XCI status	58
4.3. Non-cytological analysis for evaluating XCI status	60

CHAPTER III

Identification of the porcine *XIST* gene and its differential CpG methylation status in male and female pig cells62

1. Abstract	63
2. Introduction	65
3. Materials and Methods	69
4. Results	80
5. Discussion	104

CHAPTER IV

Some XIC-linked genes achieve dosage compensation in porcine blastocysts112

1. Abstract	113
2. Introduction	115
3. Materials and Methods	118
4. Results	129
5. Discussion	151

CHAPTER V

Identification and differential expression patterns of porcine *OCT4* variants161

1. Abstract	162
2. Introduction	164
3. Materials and Methods	167
4. Results	184
5. Discussion	212

CHAPTER VI

The effect of *OCT4* on expression of XIC-linked genes in porcine blastocysts219

1. Abstract	220
2. Introduction	222
3. Materials and Methods	225
4. Results	239
5. Discussion	252

CHAPTER VII

Correlative comparison for expression levels between pluripotent genes and XIC-linked genes in porcine blastocyst

.....	257
1. Abstract	258
2. Introduction	259
3. Materials and Methods	262
4. Results	266
5. Discussion	269

CHAPTER VIII

<i>Conclusion</i>	272
-------------------------	------------

References	279
-------------------------	------------

Abstract in Korean	311
---------------------------------	------------

LIST OF TABLES

Table III.1. The List of primer pair sequence for reverse transcription PCR	75
Table III.2. List of restriction enzymes used for each amplicon and the predicted sizes of the digested fragments	77
Table III.3. List of primer pair sequences for 5' - and 3' - RACE-PCR	78
Table III.4. List of primer pairs for bisulfite sequencing	79
Table III.5. Exons and introns information of porcine <i>XIST</i>	91
Table IV.1. Primer list used for RT-PCR	125
Table IV.2. Primer pairs used in Real-time PCR	126
Table IV.3. Primer pairs used in bisulfite-treated gDNA amplification	127
Table V.1. Primer sequence used for RT-PCR	173
Table V.2. Amplification condition for RT-PCR	174
Table V.3. Primer sequences used for real time PCR	178
Table V.4. Porcine <i>OCT4</i> variant amplification plot information	179
Table V.5. Candidate OCT4B protein isoform predicted by transcript translation	200

Table VI.1. Primer pairs used in Real-time PCR	229
Table VI.2. <i>pSIN-IRES</i> and <i>hGAPDH</i> information for linear-regression plot	236
Table VI.3. Effect of slit ZP on parthenotes development	243
Table VI.4. Effect of ZP-removal on cleaved parthenotes development	243
Table VI.5. Developmental competence of embryos with slit ZP (SZ) cultured with lentivirus	245
Table VI.6. Developmental competence of ZP-removed cleaved embryos (ZR3) cultured with lentivirus	245
Table VII.1. Primer pairs used in quantitative RT-PCR	265

LIST OF FIGURES

Figure II.1. Processes of XCI and differential timing of <i>XIST</i> expression in eutherian embryo development in different species	16
Figure II.2. Genomic composition of XIC and its homologue region in non-eutherian vertebrates	29
Figure II.3. Molecular networking of XCI-regulators, XIC-linked genes and pluripotent factors, in mouse XCI	46
Figure II.4. Chromatin modification and inactivation process of X-chromosome in differentiating mouse ES cells	54
Figure III.1. <i>XIST/Xist</i> homologue analysis of the pig genome by BLAST search	82
Figure III.2. Diagram of the candidate <i>XIST</i> expressing region in the pig	84
Figure III.3. RT-PCR analysis of each candidate region of <i>XIST</i> PCR amplicon expression	86
Figure III.4. Identified gap sequence in the pig X-chromosome scaffold, NW_003612825.1	89
Figure III.5. Identification of transcription start sites and the termination site of porcine <i>XIST</i>	93
Figure III.6. Analysis of repeat sequences in porcine <i>XIST</i> RNA	96
Figure III.7. Alignment of Illumina reads to the porcine <i>XIST</i> gene model	99

Figure III.8. Diagram of the CpG sites located in the ± 2 kb region of the transcription start site of porcine <i>XIST</i> and the methylation status of <i>XIST</i> CpG sites in male and female porcine embryonic fibroblasts	102
Figure IV.1. Identification of the porcine X-chromosome inactivation center (XIC) and sequence homology between human XIC-linked genes and their counterparts in pig	130
Figure IV.2. Mouse XIC and sequence comparison between porcine and mouse XIC-linked genes	132
Figure IV.3. Sequence homology of the human and mouse XIC-linked gene locus with counterparts on the porcine genome	133
Figure IV.4. XIC-linked gene expression in porcine embryonic fibroblasts (PEFs)	136
Figure IV.5 Porcine XIC-linked gene expression at the embryonic stage	139
Figure IV.6. Quantitative comparison of XIC-linked genes expression in male and female blastocysts	141
Figure IV.7. Diagram of CpG sites of genes in the porcine XIC	145
Figure IV.8. Methylation status of XIC-linked genes CpG sites in porcine embryonic fibroblasts (PEFs) and parthenogenic blastocysts	147
Figure IV.9. Quantitative comparison of expression levels of XIC-linked genes in porcine parthenogenic and fertilized blastocysts	149
Figure IV.10. Epigenetic dynamics of XCI status and XIC-linked gene expression in porcine developing embryos and fibroblast	160

Figure V.1. Amplification plot of each primer pair used for real time PCR	180
Figure V.2. Identification of candidate coding regions of porcine <i>OCT4</i> variants by searching for regions homologous to human <i>OCT4</i>	186
Figure V.3. Identification of porcine <i>OCT4</i> variants, <i>OCT4B</i> and <i>OCT4B1</i>	187
Figure V.4. Long and short BF3-RE5 PCR product sequence alignments to the porcine <i>OCT4</i> gene	189
Figure V.5. Porcine <i>OCT4A</i> mRNA sequence alignment to <i>OCT4A</i> homologue sequences	193
Figure V.6. Specificity validation of porcine <i>OCT4A</i> -specific primer by gDNA amplification and sequencing	194
Figure V.7. Expression of porcine <i>OCT4</i> variants in somatic tissue	198
Figure V.8. Quantitative comparison of mRNA of each porcine <i>OCT4</i> variant in mature oocytes and developing preimplantation embryos	201
Figure V.9. OCT4 variant protein expression in oocytes and preimplantation embryos	205
Figure V.10. Immunostained blastocyst without primary antibody	207
Figure V.11. Pluripotent gene expression comparison of two groups of individual blastocyst classified by expression level of each variant	208
Figure V.12. Correlative expression analysis for <i>OCT4B1</i> and pluripotent genes in individual blastocyst	210

Figure V.13. Correlation analysis between <i>OCT4</i> variant and pluripotent genes	211
Figure VI.1. Linear-regression plot for calculating copy number of transgene and integration efficiency of lentivirus stock	237
Figure VI.2. Strategies for lentivirus transduction in developing embryos	238
Figure VI.3. Correlative analysis of <i>OCT4</i> expression with XIC-linked genes in blastocyst	240
Figure VI.4. Generation of parthenogenic blastocyst with slit or removed ZP ..	244
Figure VI.5. Validation of lentiviral infection to porcine embryos	246
Figure VI.6. Effect of Human <i>OCT4</i> on parthenogenic blastocyst viral infected at cleaved embryos	249
Figure VI.7. Effect of Human <i>OCT4</i> on parthenogenic blastocyst viral infected at early blastocyst	251
Figure VII.1. Expression levels of pluripotent genes in male and female blastocyst	267
Figure VII.2. Correlative expression analysis between pluripotent genes and XIC-linked genes	268
Figure VIII. XCI in porcine embryo development	278

ABBREVIATES

ANOVA	Analysis of variance
BLAST	Basic Local Alignment Search Tool
BSA	Bovine Serum Albumin
cDNA	Complementary DNA
CHIP	CHromatin-ImmunoPrecipitation
COC	Cumulus-Oocyte Complex
CTD	C-terminal Transactivation Domain
DMEM	Dulbecco's Modified Eagle Medium
DMR	Differentially Methylated Region
DPC	Days Post Coitus
DW	Distilled water
EC	Embryonic Carcinoma
eCG	Equine Chorionic Gonadotrophin
EGF	Epidermal Growth Factor
EGFP	Enhanced Green Fluorescent Protein
EGFP-lentivirus	EGFP-integrating lentivirus
EpiSCs	Epiblast Stem Cells
ES	Embryonic stem
EtBr	Ethidiobromide

FBS	Fetal Bovine Serum
FISH	Fluorescent In Situ Hybridization
gDNA	Genomic DNA
hCG	Human Chorionic Gonadotrophin
hGAPDH	Human GAPDH
hOCT4	Human OCT4
hOCT4-lentivirus	hOCT4-integrating lentivirus
HPC	Hours post coitus
HRP	Horseradish Peroxide
ICM	Inner Cell Mass
iPSCs	Induced Pluripotent Stem Cells
IRES	Internal Ribosome Entry Site
LIF	Leukemia Inhibitor Factor
LNAs	Locked Nucleic Acids
mRNA	Messenger RNA
MSCI	Meiotic Sex Chromosome Inactivation
mTBM	Modified Tris-Buffered Medium
MYA	Millions years ago
ncRNA	non-coding RNA
NTD	N-terminal Transactivation Domain
ORF	Open Reading Frame
PCR	Polymerase Chain Reaction
PEF	Porcine Embryonic Fibroblast

PEG	polyethylene glycol
pFF	Porcine Follicular Fluid
POU5F1	POU domain, class 5, transcription factor 1
PRC	Polycomb Repressive Complex
PURO	Puromycine registence gene
PZM-3	Porcine Zygote Medium 3
RACE PCR	Rapid Amplification of CDNA Ends
RNAi	RNA interference
RT-PCR	Reverse Transcription-Polymerase Chain reaction
S.E.M	Standard Error Mean
SCNT	Somatic Cell Nuclear Transfer
SRA	Sequence Read Archive
SZ	Slit ZP
TCM	Tissue Culture Medium
TSS	Transcription Starting Site
UTR	Untranslated region
XCI	X-Chromosome Inactivation
XIC	X-chromosome Inactivation Center
XIST	X-chromosome Inactivation-Specific Transcript
Xm	Maternal X-chromosome
Xp	Paternal X-chromosome
ZGA	Zygotic genome activation
ZP	Zona Pellucida

ZR3	ZP-removed at embryonic day 3
ZR5	ZP-removed at embryonic day 5

CHAPTER I

General Introduction

General Introduction

Dosage Compensation is essential process for balancing the expression levels of genes in sex chromosomes between males and females. There are three typical mechanisms for equalizing dosage of genes located in sex chromosome. Genes in one X-chromosome transcribed two-fold high in male flies (Park and Kuroda 2001) and both X-chromosomes are partially inactivated in worms (Meyer 1999). Contrary to the non-vertebrates species, placental mammals apply inactivation of one X-chromosome in females and balancing the dosage of X-linked genes between males and females (Lyon 1961; Okamoto and Heard 2009). The epigenetic process for matching the expression levels of X-linked genes between male and female eutherians is called X-chromosome inactivation (XCI). After the first suggestion of the event by observing condensed chromatin structure in nucleus called 'barr-body' (Lyon 1961), numerous studies have been performed and help to understand the event comprehensively. Fifty years on, exact XCI mechanism in developing embryos has been confirmed and various factors have been reported to be associated with XCI in mice.

One report observed that translocation of the specific genomic region in X-chromosome to autosome induced *cis* inactivation of the autosome. This study

suggested that the fragment in X-chromosome would trigger chromosomal silencing and named X-chromosome inactivation center (XIC) (Rastan 1983) and led to identification of X-chromosome inactivation specific transcript (*XIST/Xist*) (Brockdorff *et al.* 1991; Brown *et al.* 1991a; Brown 1991) as a key regulator for *cis* inactivation of X-chromosome (Penny *et al.* 1996). This non-coding RNA (ncRNA) gene showed similar genomic structure with the large first and last exons, and some small exons between two large exons in human and mice (Brockdorff *et al.* 1992; Brown *et al.* 1992). The gene has been also reported to be originated by disruption of ancestral protein coding gene, *Ln3*, and integration of mobile element (Duret *et al.* 2006; Elisaphenko *et al.* 2008). The evolutionary short period for evolving the gene resulted in less conserved sequence of the ncRNA despite its genomic structural similarity among the eutherians (Romito and Rougeulle 2011). This genetic property observed in this gene makes hard identify the orthologs in other species and the gene has been identified only in a few species including human and mouse (Chureau *et al.* 2002; Hwang *et al.* 2013a) even though the gene is the most important factor for XCI.

Some studies applied *XIST* expression for judging normality in embryo development and status of pluripotent stem cells (Nichols and Smith 2009; Inoue *et al.* 2010). As *XIST* is master gene for inactivation of X-chromosome and the process is directly related to preimplantation embryo development, abnormal

expressions of the gene was observed in embryos of various species produced *in vitro* (Inoue *et al.* 2010; Matoba *et al.* 2011). For example, cloned mouse embryos showed abnormal expression of *Xist* and skewed XCI (Inoue *et al.* 2010) but deletion of the one *Xist* allele or RNA interference (RNAi)-mediated repression of the gene rescue the birth rate of cloned mice (Matoba *et al.* 2011). Recent report in embryonic stem (ES) cells showed that the inactive status of X-chromosome is different following the pluripotent status (naïve vs primed) because the origin of these embryonic cells are different each other (Brons *et al.* 2007; Tesar *et al.* 2007; Nichols and Smith 2009). This feature suggested the *XIST* expression or methylation status of its promoter region is represent markers for elucidating pluripotency of ES cells. These studies showed the importance of the *XIST* genes in various research fields.

Although the first suggestion of XIC by Rastan in 1983 led to identification of *XIST*, the importance of the genes located in the region has been highlighted recently. This genomic center has been undergone dynamic evolution and it is confirmed by presence of XIC-homologue region in non-eutherian vertebrates (Romito and Rougeulle 2011). The most represent difference of the synteny between eutherian and non-eutherian vertebrates is presence of ncRNAs including *XIST* (Hore *et al.* 2007). Although conserved protein coding genes, like *CHIC1*, *SLC16A2*, and *RLIM* are coded on the close genomic distance in vertebrates, the

orthologs of five non-eutherian vertebrate protein-coding genes located in the internal region between *CHIC1* and *SLC16A2* were not discovered in eutherians but they showed partial sequence homology with the ncRNAs in the counterpart region in eutherians. This evolutionary break-point has been considered to occur during adaptive radiation of mammal as the synteny was disappeared in marsupials. Newly arising ncRNAs in eutherian, like *Tsix*, *Jpx*, and *Ftx*, have been reported to regulate *XIST* expression and XCI in mouse (Lee and Lu 1999; Jonkers *et al.* 2009; Tian *et al.* 2010). However, the knowledge on this evolutionary diverse genomic region among the species is little in other species despite the important relations with the XCI and ncRNAs in XIC.

Two waves of XCI exist in developing mouse embryos. The first wave is imprinted inactivation of paternal X-chromosome (Xp) (Mak *et al.* 2004; Okamoto *et al.* 2004). This process is observed in early cleaved embryos (4-8 cell stage) and induced by *Xist* expression from paternal allele (Kay *et al.* 1994; Mak *et al.* 2004; Okamoto *et al.* 2004). *Cis* acting of transcribed paternal *Xist* let the Xp inactive. The inactive Xp is reactivated in inner cell mass (ICM), and followed by inactivation of one X-chromosome randomly (Wutz and Jaenisch 2000). This second wave is observed in ICM but trophoblast, which is other compartment lineage in blastocyst, maintains the imprinted Xp inactivation. Contrary to XCI in mouse embryos, recent one study suggested that human and rabbit have their own

inactivation process of X-chromosome during embryonic stages differing to that of mice (Okamoto *et al.* 2011). Represent two differences on XCI process between mouse and non-mouse species are 1) presence of imprinting XCI and 2) late time window for initiating XCI in human and rabbit embryos. Human and rabbits showed absence of imprinting XCI and early blastocysts of human didn't accomplished inactivation of X-chromosome even though accumulated *XIST* transcript was observed biallelically in both ICM and trophoblast. This study strongly suggested that the XCI event occurs with species-specific manner. Although the orthologs of *XIST* didn't discovered in marsupials which is other placental mammals diverged about 150 - 180 million years ago with eutherians from ancestral mammal, they are considered to recruit imprinted XCI as strategy for dosage compensation of X-linked genes by meiotic sex chromosome inactivation (MSCI) in paternal germ line (Namekawa *et al.* 2007). This would mean the imprinted XCI in mouse embryo is remaining properties of ancestral XCI even though there are differences on XCI in embryonic development among the placental mammals. These differences of XCI among the placental mammals showed that this epigenetic process for compensating the dosages of sex chromosome is diverse among the species and highlight the requirement of the species-specific studies on the epigenetic event. However, only less number of studies has been carried out in non-mouse eutherians.

As described above, various factors are reported to be associated with XCI and *Xist* expression in mouse. Various genes and genomic element in X-chromosome is tightly linked with XCI and autosomal genes like pluripotent factors has been suggested to be related with *Xist* expression (Navarro *et al.* 2008; Donohoe *et al.* 2009; Ma *et al.* 2011; Gontan *et al.* 2012). Represent pluripotent factors, like Oct4, Sox2, and Nanog has been reported to bind to the first intron of *Xist* and considered to repress *Xist* expression in undifferentiated mouse ES cells (Navarro *et al.* 2008). Other factors related to pluripotency like *Rex1*, *Klf4*, and *Myc* were also reported to regulate the *Xist* antisense transcript, *Tsix* (Navarro *et al.* 2010; Payer *et al.* 2013). These reports suggest that pluripotent circuits are closely linked with the XCI although one report observed that deletion of the intronic region of *Xist* gene didn't affect to XCI in undifferentiated mouse ES cells (Minkovsky *et al.* 2013). However, the relevance between XCI and pluripotent factors is still unclear in other species.

Numerous studies have been conducted to understand this complex epigenetic event after first suggestion of the XCI for over fifty years using mouse model. However, recently studies strongly addressed that the process is evolutionary diverse mechanism among the eutherians. Therefore, studies regarding XCI need to be carried out in various eutherians for comprehensive understanding of this mechanism. So in this study, the knowledge about XCI was tried to be established in pigs which has been suggested to be a model animal for biomedical and

developmental science. For the insight about XCI in pigs, its main inducer, *XIST*, and gene cluster regulating the process, XIC, were identified. And epigenetic changes of the XIC-linked genes were examined in preimplantation embryos to know the XCI process during porcine embryo development. After confirming the process in porcine preimplantation embryos, regulations of XCI by pluripotent factors were confirmed.

CHAPTER II

Literature review

Literature review

1. X-chromosome inactivation in mammals

X-chromosome inactivation (XCI) is an epigenetic process occurred in placental mammals. The event induces silencing of genes in one of X-chromosomes to equalize expression levels of the genes between male and female adults. Although various studies have been carried out to confirm and understand the process of XCI after its first observation (Lyon 1961), the large proportion of knowledge regarding this epigenetic event is restricted in mouse models. However, this process is achieved with differential manners among the placental mammals and even in eutherians during their embryonic developmental period (Okamoto *et al.* 2011). For example, inactivation of X-chromosome in ancestral placental mammals (marsupials) has been considered to be achieved by passing paternal inactive X-chromosome and the imprinted XCI pattern would be sustained during their full-term development (Namekawa *et al.* 2007). Imprinted XCI is also observed in early stage of mouse preimplantation embryos and their extraembryonic tissue development (Augui *et al.* 2011). However, eutherians, including mice, showed random XCI in their embryonic part and imprinted XCI in earlier stage of embryos were not observed in rabbit and human embryos (Okamoto *et al.* 2011). Therefore, XCI is an evolutionary dynamic process (Okamoto and Heard 2009; Escamilla-

Del-Arenal *et al.* 2011) and detailed explanation of the event in each mammals is described below.

1.1. X-chromosome inactivation in monotremes and marsupials

Monotremes and marsupials are ancient mammals which are evolutionary diverged about 210 million years ago (MYA) with marsupials and 180 MYA with eutherians, respectively. Monotremes like platypus have five X-chromosomes and the chromosome showed homology with sex-chromosome in birds (Z) and therians (X) (Grutzner *et al.* 2004). This differential number of sex chromosomes in monotremes with marsupials and eutherians indicates that the egg-laying animals are evolutionary break point for evolving X and Y chromosomes. Interestingly, this animal wouldn't require the X-chromosome inactivation center (XIC) for achieving dosage compensation by inactivating X-chromosome because the sequence homologue region is detected in autosome (chromosome 6), not in X-chromosomes (Hore *et al.* 2007). Indeed, the dosage compensation of the X-linked genes are partially achieved in adult and considered to occur locus-specifically (Deakin *et al.* 2008). This birds-like compensation status is thought to be reasoned by stochastic inactivation of X-borne genes and this is suggested to be remained ancient traits for dosage compensation (Deakin *et al.* 2009).

Compared to monotremes, marsupials like wallaby and kangaroo showed mature

sex chromosome as eutherians. These animals are reported to apply imprinted XCI for dosage compensation because paternal X-chromosome (Xp) preferential inactivation was observed in adult tissues of marsupials (Richardson *et al.* 1971). Other study observing the late DNA replication of Xp in marsupials also support the Xp-preferential inactivation (Sharman 1971). This mechanism was considered to be induced by delivering inactive X-chromosome from paternal haploid (Namekawa *et al.* 2007). Contrary to the XIC-homologue region in monotremes, its counterpart was observed in the X-chromosome. However, the genomic region is disrupted and *LNK3* which is ancestral protein coding gene and observed in amphibians, reptiles, and birds but not in eutherians is present instead of *XIST* which is reported to be evolved from *LNK3* and mobile elements (Duret *et al.* 2006). This means marsupials don't have *XIST*-mediated XCI for compensating genes in X-chromosome. However, one study raised the possibility that the animals would apply the XCI mechanism similar to eutherians by using non-coding RNAs (ncRNAs) (Grant *et al.* 2012). X-linked ncRNA called *RSX* expressed female specifically and coat the inactive X-chromosome as *XIST*-cloud observed in inactive eutherian X-chromosome. Although this 25 Kb of transcript has no sequence homology with *XIST*, *RSX* transgenes in autosome of mouse embryonic stem (ES) cells showed *cis* inactivation of the chromosome. This result demonstrates that the gene is functional counterparts of *XIST* in eutherians and highlighted the possibility that marsupials would have process synonymous with *XIST*-dependent XCI as eutherians. The inactive status of X-chromosome in

marsupials is considered to be unstable compared to the eutherians because the CpG sites of inactive X-chromosome was not methylated and transient reactivation of inactive X was observed in cultured marsupial cells (Kaslow and Migeon 1987; Hornecker *et al.* 2007).

1.2. X-chromosome inactivation in mice

Large number of studies defined epigenetic changes of X-chromosome in mouse embryo development. Briefly, two waves of XCI, imprinted and random XCI, are appeared during embryo development (Figure II.1). Imprinted XCI is observed in early preimplantation embryos and extraembryonic lineage. Imprinted inactive Xp is removed and reactivated in inner cell mass (ICM). After that, one of the active X-chromosome in ICM is subject to be inactivated randomly and maintained the mark in adults. Each step of XCI, imprinted XCI, reactivation of Xp, and random XCI will be described below.

Imprinted XCI Imprinted XCI is observed in mouse preimplantation embryos and extraembryonic tissues. Although the reason why Xp is chosen for the transient inactivation, numerous studies revealed the process of the imprinted XCI. The Xp-preferential inactivation had been thought to be originated by delivery of Xp which is inactivated during spermatogenesis (Huynh and Lee 2003). The study observed

that silenced Xp in 2-cell embryos, therefore, suggested that the inactive Xp would be resultant of MSCI which would be reasoned by asynapsis of sex chromosome in male gremlin (Turner *et al.* 2006). However, it was reexamined and confirmed that transient activation of Xp is present in mouse 2-cell embryos (Patrat *et al.* 2009; Namekawa *et al.* 2010). And also, XIC integrated in paternal autosome showed imprinted inactivation in female placenta (Okamoto *et al.* 2005). This result means MSCI-mediated inactivation was not happened in inactivated autosome harboring XIC. Therefore, the imprinted XCI is considered to be carried out after temporary activation of silenced Xp and MSCI would have less effect on the imprinted XCI in mouse preimplantation embryos. *Xist* expression is initiated at 2-cell stage from Xp and Xp-linked genes activated by zygotic genome activation (ZGA) were silenced rapidly, and rapid chromosome-wide inactivation was detected at the second division of embryos (Okamoto *et al.* 2004). Contrary to Xp, maternal X (Xm) maintains the transcriptionally active status by expression of *Tsix* repressing *Xist* (Sado *et al.* 2001). And also, another *trans* acting factor, *Rlim*, was reported to induce imprinted XCI (Shin *et al.* 2010). Although *Xist* is main factor for inactivating one X-chromosome and equalizing dosages of X-chromosome between males and females, its requirement in imprinted XCI during preimplantation embryo development is controversial. Two studies carried out by different groups showed depletion of *Xist* was or was not induced imprinted XCI in preimplantation embryos (Kalantry *et al.* 2009; Namekawa *et al.* 2010). However, the importance of gene for sustaining imprinted XCI in extraembryonic lineage is

clear because postimplantation embryos without paternal *Xist* failed to develop and form mature placenta (Marahrens *et al.* 1997). Xp inactivated by imprinting showed differential properties on its maintenance compared to the inactive X-chromosome selected randomly. Contrary to randomly inactivated X-chromosome, imprinted silencing of Xp needs persistent *Xist* expression (Ohhata *et al.* 2011) and Eed, which is subunit of histone methyltransferase, polycomb repressive complex (PRC) 2, (Wang *et al.* 2001), but is independent to DNA methylations on promoter of inactive X-linked genes (Sado *et al.* 2000). These properties are similar to the randomly inactivated X-chromosome observed in ES cells 2 - 3 days after differentiation (Escamilla-Del-Arenal *et al.* 2011). Therefore, chromosomal status of imprinted inactive X-chromosome might be less stable compared to the X-chromosome inactivated randomly.

Figure II.1. Processes of XCI and differential timing of *XIST* expression in eutherian embryo development in different species. In mice, imprinted XCI is observed on paternal X-chromosome (Xp) during cleaved embryos after zygotic genome activation (ZGA). This memory is disappeared and inactivated Xp is reactivated at the ICM of late blastocyst. However, placental lineage maintains imprinted XCI during full-term development. In humans and rabbits, however, imprinted XCI is not observed and generally considered to be applying random XCI for their dosage compensation. Both chromosomes in late human blastocyst are active although *XIST* is biallelically expressed and coat both X-chromosomes. In rabbit, some cell populations in trophoctodermal cells showed both inactive X-chromosomes in early blastocyst but one inactive X-chromosome is observed in ICM and trophoctodermal cells after developed to late stage of blastocyst. In cows and pigs, it is unclear that allele-dependent expression of *XIST* and presence of imprinted XCI in early stage of embryo development. However, it was considered that dosage compensation in embryonic day 7 bovine blastocysts wouldn't be completed chromosome-widely and the proportion of compensated X-linked genes is increased gradually following embryo development. And XCI in placental lineage is thought to be achieved by imprinting. In pigs, *XIST* was considered to be expressed maternal X-chromosome preferentially and inactivation of X-chromosome would be completed in day 15 elongated blastocyst or earlier than the developmental stage. Black arrows under the diagram representing XCI and *XIST* expression of developing embryo of each species indicate *XIST* positive stages.

Reactivation of Xp Inactive Xp is reactivated in ICM of blastocysts (E4.5). This epigenetic change was achieved by repression of *Xist* (Sheardown *et al.* 1997a), removal of heterochromatin mark, H3K27ME3, and subunits of PRC2, which is methyltransferase to H3K27, Ezh12 and Eed (Okamoto *et al.* 2004). Main factor inducing chromatin change is *Tsix*, *Xist*-antisense transcript (Lee *et al.* 1999). Transient expression of *Tsix* in E4.5 blastocyst could reactivate imprinted Xp but inactivated Xp in extraembryonic lineages failed to maintain the active status and returned to transcriptionally silenced status (Ohhata *et al.* 2011). This result indicates that even though *Tsix* is critical for reactivating XCI by repressing *Xist* (Sado *et al.* 2001; Sado *et al.* 2005), other factors like pluripotent factors which are expressed in ICM and epiblast preferentially would be associated with the epigenetic changes of inactivated Xp. Indeed, several studies have been revealed that pluripotent factors could repress *Xist* (Navarro *et al.* 2008; Donohoe *et al.* 2009; Ma *et al.* 2011; Gontan *et al.* 2012) and activate *Tsix* (Navarro *et al.* 2010; Payer *et al.* 2013). Detailed explanation of the molecular networking will be described later. This reversible property of inactivated Xp achieved by imprinting might be caused by less stable inactive status of the chromosome (Ohhata and Wutz 2013).

Random XCI One of the X-chromosome is inactivated in epiblast after reactivation of imprinted X-chromosome, subsequently (Mak *et al.* 2004). In this

second wave of XCI, future inactive X-chromosome is chosen randomly. This is firstly discovered by observing maternal *Xist* expression in epiblast (E5.5) (Mak *et al.* 2004). The chromatin status of inactivated X-chromosome is more stable and compact compared to the inactivated Xp by imprinting as described above. Randomly inactivated X-chromosome in epiblast shows sensitivity to some epigenetic modifiers like *SmcHD1* and *Dnmt1* to maintain inactive status of X-chromosome (Sado *et al.* 2000; Blewitt *et al.* 2008). Indeed, this inactive X-chromosome showed methylated promoters of X-linked genes and this is considered to make the chromosome irreversibly inactive. As similar to the silencing of X-chromosome in ES cell differentiation, which required several days for stabilizing inactive X-chromosome thoroughly (Escamilla-Del-Arenal *et al.* 2011), it is thought that stable inactivation of one X-chromosome in epiblast would be demanded some days. Silencing of some genes in X-chromosome, like *Hprt1* and *Pgk1*, were observed at the E6.5 - 7.0 embryos by examining their enzymatic activity (Monk and Harper 1979; McMahon and Monk 1983). However, the methylation of *Hprt1* promoter on inactive X-chromosome, which is indicator for stabilized inactive status, is confirmed after about one week later (E13.5) (Lock *et al.* 1987). Interestingly, promoter methylation of *Pgk1* was observed rapidly (E7.0) compared to that of *Hprt1* (Singer-Sam *et al.* 1990) and this might mean stable inactivation of X-linked genes is achieved gradually in random XCI.

1.3. X-chromosome inactivation in non-mouse eutherians

Compared to the studies in mice, only restricted information regarding XCI in non-mouse eutherians has been reported. However, the studies in the species revealed differential time window and strategies for XCI compared to the mouse XIC. In here, XCI in four eutherians, human, rabbits, cows, and pigs, will be described and highlight the diversities of the mechanism even among the eutherians (Figure II.1).

Human The first study examining *XIST* expression in human preimplantation embryos showed that the gene was not expressed sex-dependently (Daniels *et al.* 1997). Although the study was conducted about 20 years ago, detail process of XCI in human preimplantation embryos was suggested recently (Okamoto *et al.* 2011). As the study reported in 1997, *XIST* expression was observed in both sex of morulae and blastocysts, and this means imprinted *XIST* isn't expressed at least in human preimplantation embryos. Add to the difference of the process between human and mouse embryos, the authors reported that *XIST* is expressed biallelically by observing the two *XIST* cloud in female blastocyst. However, although the X-chromosome was coated with *XIST*, examined three X-linked genes, *ATRX*, *FGD1*, and *HUWE1*, were still expressed on the X-chromosome. Also, one of the epigenetically repressive chromatin markers, H3K27ME3, was not detected on the *XIST*-coated X-chromosome. This result indicates chromosome-wide XCI in

human preimplantation is initiated after blastocyst stage and this time window is later than in mice. And the result that there is no difference on expression patterns of *XIST* and examined three genes between human ICM and trophoblast demonstrated that human blastocyst doesn't employ the distinct mechanism following the lineages. These result also supports that humans have distinct mechanism for inactivation of X-chromosome in preimplantation embryos. However, one study showed differential results with those of Okamoto and his colleagues' (van den Berg *et al.* 2009). They reported that *XIST* foci were observed at 8-cell stage first and monoallelic *XIST*-coating was mainly distributed at the blastocyst. Add to here, examined X-linked genes (*CHIC1* and *COT1*) in blastocyst showed their primary transcription on the *XIST*-non-coated X-chromosome, therefore they addressed XCI is accomplished in blastocyst. The differences on timing for *XIST* expression and number of X-chromosome with *XIST*-cloud in blastocyst would be reason by technical differences. *CHIC1* of which expression was examined in van den Berg and colleagues' study has been known to be silenced in 4- to 8-cell mouse embryos (Patrat *et al.* 2009). And also, as the timing that one allele of X-borne genes is inactivated is different among the genes (Kalantry *et al.* 2009), these differential result on the silencing of one allele of X-linked genes would be reasoned by the differences on the examined genes. Although some differential results were reported in two studies, it is clear that the mechanism in early embryo is different between human and mice. XCI in human extraembryonic tissues had been controversial because it is unclear whether the

placental lineage cells use imprinted XCI as in mouse. Early studies suggested that XCI in human extraembryonic tissues is induced through imprinting by observing inactive Xp in cytotrophoblast (Harrison 1989) and trophoblast (Goto *et al.* 1997). However, recent report refuted the previous results and clarified that the XCI in human placenta occurred randomly by examining allele-specific expression of X-linked genes (Moreira de Mello *et al.* 2010).

Rabbit XCI in rabbits has not been studied much, but one report clearly demonstrated the XCI in rabbits is different with that in mouse embryos (Okamoto *et al.* 2011). First *XIST* expression was detected at 8-cell stage embryos and biallelic expression was observed until early blastocysts. Interestingly, some trophoblast populations in early blastocyst (hours post coitus (HPC) = 96) showed *XIST*-coated two X-chromosomes and silenced *HPRT1* expressions. However, inactivation of both X-chromosome in the trophoblasts are diminished after developed to late blastocyst and the most cells in the embryos showed inactivation of one X-chromosomes. Compared to the trophoblast, ICM showed few cell populations with *XIST*-accumulated X-chromosome in early blastocyst and progression of XCI in one X-chromosome was observed in the ICM of late blastocyst. It looks like that rabbit XCI undergone two waves of inactivation of X-chromosome and the species doesn't accompany the imprinted XCI because blastocyst generated by parthenogenesis didn't showed differences on the XCI patterns compared to those

of fertilized embryos. Therefore, it is concluded that rabbits have developmentally unique XCI process and its time window for initiation is later than mice.

Cow Early studies in bovine preimplantation embryos revealed that obvious expression of *XIST* is detectable after 8-cell stages (De La Fuente *et al.* 1999). After the study, some studies focused on the sexual dimorphic expressions of X-linked genes in embryonic stages (Gutierrez-Adan *et al.* 2000; Peippo *et al.* 2002; Wrenzycki *et al.* 2002). These PCR-based studies suggested that *XIST* expressed in male and female preimplantation embryos (after morular stage) and sexual dimorphism on expressions of X-linked genes was observed in blastocyst. This means that dosage compensation of X-linked genes would not be completed until blastocysts in cows. And the possibility was supported by recent one study examining differential expression levels of genes between male and female bovine blastocysts by using microarray (Bermejo-Alvarez *et al.* 2010). The study revealed that most of X-linked genes of which expression is detected in blastocysts showed female preferential expression. This sexual dimorphic expression in blastocyst indicates dosage compensation of the X-borne genes is not accomplished chromosome-widely. And as about 70% of female preferentially expressed X-linked genes expressed less than 1.66-fold high compared to the male blastocysts, it was considered that XCI would be initiated in some cell populations in this embryonic stage. Indeed, elongated bovine embryos which are developmentally

late stage compared to blastocyst and considered to start gastrulation showed equalized expression levels of five X-linked genes among the examined seven genes which were expressed female preferentially in blastocysts (Bermejo-Alvarez *et al.* 2010; Bermejo-Alvarez *et al.* 2011). These results support that XCI in bovine blastocyst is not accomplished and the species have late timing for chromosome-wide inactivation of X. However, it is unclear whether they applying imprinted XCI during the early embryonic stages. Some reports suggested that bovine placenta exhibit inactive Xp (Xue *et al.* 2002) and paternal *XIST* expression (Dindot *et al.* 2004b). However, it is still unclear the initiation of XCI in early developing embryo is affected by imprinting. Therefore, cytological approaches for confirming exact mechanisms for achieving XCI in bovine embryos are also conducted further.

Fig Compared to the other species described above, knowledge on porcine XCI is most little and unclear. One study examined the epigenetic chromatin changes in embryos and they observed that represent heterochromatin marker, H3K27ME3, was globally detected in porcine blastocyst but its foci were observed female specifically in the epiblast and trophoblast of spherical blastocyst (Gao *et al.* 2011). And also, another recent study examined sexual dimorphic expression of some X-linked genes in male and female blastocyst originated differentially (Park *et al.* 2012). Although there were the variations of the gene expression levels in

individual blastocyst following the embryo origin (obtained by *in vivo*, *in vitro* fertilization, and cloning), examined five X-linked genes, *XIST*, *HPRT1*, *G6PD*, *BEX1*, and *PGK1*, showed female preferential expression regardless of the blastocyst origin. These genes were considered not to be the genes escaping XCI in pigs and, therefore, the result supports the possibilities that XCI is not accomplished in blastocyst stage. One report observed that *XIST* would be Xm preferentially expressed in porcine preimplantation embryos by comparing expression levels of the genes among the blastocyst produced by *in vitro* fertilization, parthenogenesis, and androgenesis (Park *et al.* 2011). However, as the report carried out by using pooled embryos without distinguishing sex of each embryos, imprinting expression of the gene in porcine embryonic stage is need to be examined again. Unfortunately, these previous reports were conducted prior to the identification of complete porcine *XIST* gene, so cytological approach for examining the XCI process in embryo was impossible. But, recently, porcine *XIST* and its regulatory regions were identified (Hwang *et al.* 2013a) and this will support the further studies on the porcine XCI.

Previous studies regarding XCI in various species obviously showed that the process is evolutionary diverse among the eutherians. Even the master gene for the process is common, previous studies demonstrated that each species employ its specific mechanism for inactivating one of X-chromosomes in females. Although

the reason for these differences is unclear, one hypothesis is based on the differential timing for ZGA among the eutherians. ZGA occur at 1- to 2- cells in mice (Bouniol *et al.* 1995), 4- to 8-cells in rabbits and humans (Manes 1973; Braude *et al.* 1988), 8- to 16- cells in bovine (Frei *et al.* 1989), and 4-cells(Jarrell *et al.* 1991). These fast timing for ZGA in mouse compared to other eutherians would result in unstable chromatin structure for expressing *Xist* in mouse Xm. Whilst, ZGA in other species occur after a few round of cell division, and it would be possible that *XIST* expressed permissively from both allele. This differences might induce biallelic expression of *XIST* in other species rather than imprinted XCI (Sado and Sakaguchi 2013). However, there are opposite opinions that imprinted XCI wouldn't be related to ZGA timing by evidencing XCI in cow (Dupont and Gribnau 2013) because the species has late ZGA timing but some studies reported imprinted XCI in bovine (Xue *et al.* 2002; Dindot *et al.* 2004b). Therefore, more studies in various species are required further to understand the mechanism comprehensively among the species.

2. Regulators in X-chromosome inactivation

Various regulators involved to XCI mechanism have been reported after first suggestion of candidate genomic region for *cis* inactivation of X-chromosome (Rastan 1983). The genomic region, called XIC, has been identified by observing translocation of certain fragment of X-chromosome make fused chromosome

inactive. This finding leads to identification of *XIST* which is key factor for initiating XCI (Brockdorff *et al.* 1991; Brown *et al.* 1991b). And recently, various factors in XIC and pluripotent factors have been reported to be associated with initiating and maintaining XCI. Detail information for each factors are described below.

2.1. X-chromosome inactivation center

The center has been reported firstly in 1983 (Rastan 1983). After identification of the region, comparative sequence comparison among the various species revealed that the genomic region is syntenic among even non-eutherian vertebrates as well as eutherians (Duret *et al.* 2006). The genes in this short genomic region (about 0.5 to 1 Mb) located specific order and the genic structure and gene composition of the region showed conservation among the species (Figure II.2). However, distinct differences between the eutherians and the other species are observed. Although the genes in each side of XIC are conserved, newly evolved ncRNAs are observed only in the eutherian XICs (Chureau *et al.* 2002; Duret *et al.* 2006). Indeed, although the counterparts of eutherian XICs are observed in non-eutherian vertebrates, they are located on autosomes and the XIC-homologue region in X-chromosome was firstly observed in the marsupials (Duret *et al.* 2006; Hore *et al.* 2007). The evolutionary changed location of XIC-homologue regions raised the possibility that the XIC-linked genes discovered only in eutherians are

closely related to XCI as the process is derived only in therians which apply differential manner for inactivating X-chromosome between marsupials and eutherians. And also, the genomic region expressing eutherian-specific XIC-linked genes are disrupted and each side of genomic region is apart in monotremes and marsupials (Grutzner *et al.* 2004; Hore *et al.* 2007). This also support that the genomic region is closely related to eutherian-specific XCI process and originated rapidly because the genomic location and compartment of the region is distinct even among mammals. Therefore, the examining of the XIC-linked genes in XCI has been focused recently and it is important to analyz the roles of each factor in various species because the regulators are also genetically less conserved among the eutherians because of its origination by rapid evolution.

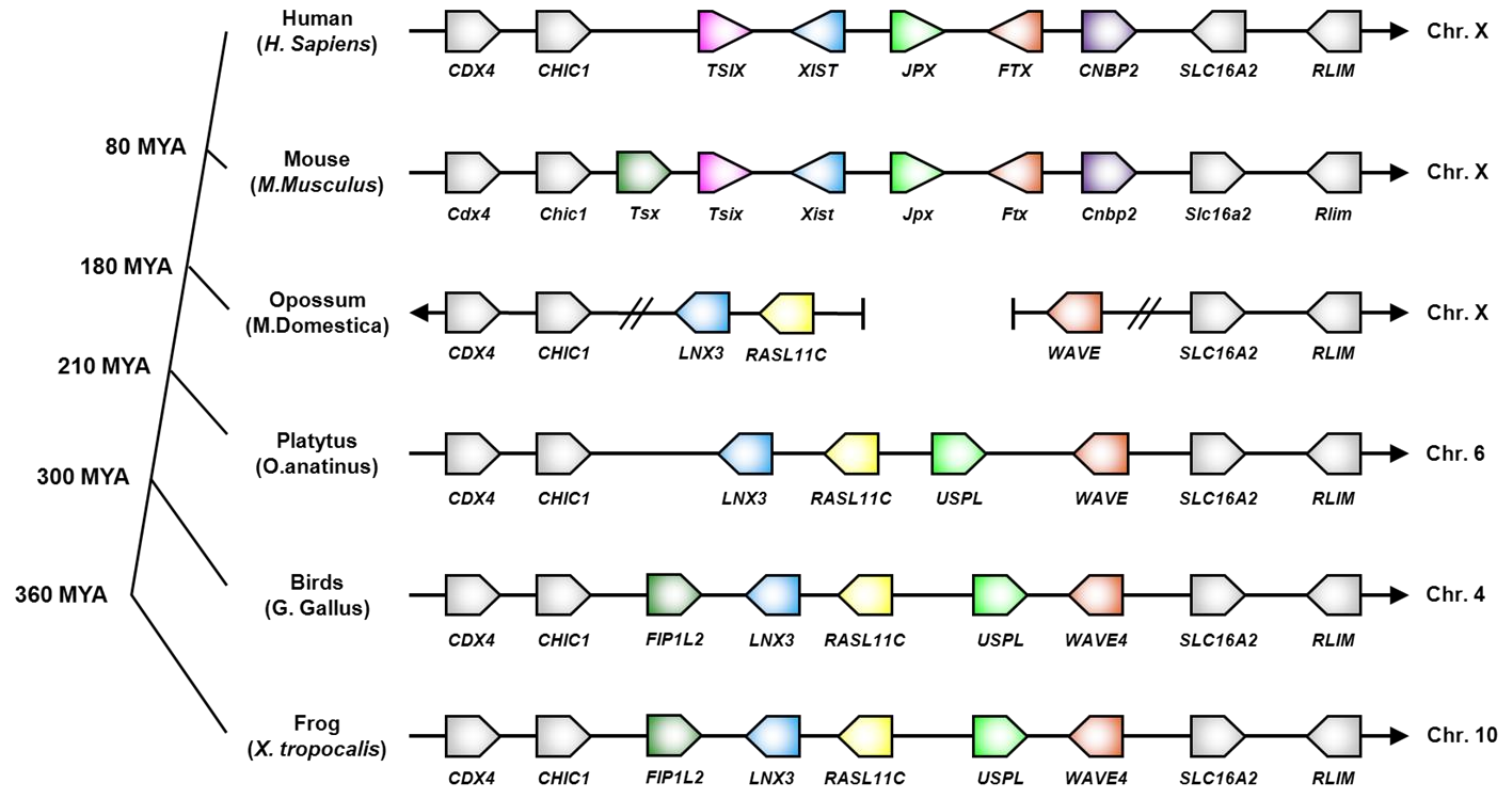


Figure II.2. Genomic composition of XIC and its homologue region in non-eutherian vertebrates. Arrows and triangles indicate protein coding and non-coding RNA (ncRNA) genes, respectively. Highly conserved protein coding genes flanking ncRNAs (*CDX4*, *CHIC1*, *SLC16A2*, and *RLIM*) is represented to gray colors. The components with same colors in each XIC and XIC-homologue region are orthologs or shared homologue region. For example, sequence homologue region to *LNX3* is observed in the human and mouse *XIST* gene. The diagram and phylogenetic tree is not scaled.

XIST This ncRNA was identified first in human by female-specific hybridization of a human cDNA probe (Brown *et al.* 1991a) and confirming the probe is located in XIC (Brown *et al.* 1991b). After identification of the orthologs in various species (Brockdorff *et al.* 1991; Chureau *et al.* 2002; Hwang *et al.* 2013a), comparative sequence analysis revealed that this gene has been originated by pseudogenization of *LNK3* genes (Duret *et al.* 2006; Elisaphenko *et al.* 2008). The orthologs of *LNK3* are not present in eutherian genomes and only partially homologue to 5' and 3' end region of *XIST* orthologs. And *LNK3* non-homologue genomic sites in *XIST* orthologs are comprised with mobile element. Various studies revealed that this polyadenylated ncRNA is main factors for the inducing XCI. The embryo inherited mutated paternal *Xist* allele in mouse showed lethality because of abnormal development of extraembryonic tissues (Marahrens *et al.* 1997; Kalantry *et al.* 2009). This means the gene is closely related to the imprinted XCI mechanism. And also, fetal cells, which are originated from epiblast and undergone random inactivation of X-chromosome, with heterozygotic mutated *Xist* allele showed preferences on the inactivation of wild-type X-chromosome (Marahrens *et al.* 1998; Kalantry *et al.* 2009). In addition, differentiating *Xist*-null ES cells were failed to inactivate X-chromosomes (Penny *et al.* 1996). These results clearly support the essential functions of *XIST* on the XCI. This gene consists of several repeat sequence regions (repeat A to F) and among them, repeat A has been reported to essential compartment for silencing the X-chromosome (Wutz *et al.* 2002). This 24 bp of monomer is repeated 7.5 times in mouse *Xist* and form secondary structure of

double hair pin. Interestingly, this region expresses short transcript independent to *Xist* (Rep A) and this short transcript is reported to recruit PRC2, which induce repressed chromatin status by modifying histone (Zhao *et al.* 2008). Another repeat region, repeat C, which is dominantly present in mouse *Xist* (14 copies) but less popular in human *XIST* (1 copy) is reported to associated with localization of *XIST* to future inactive X-chromosome (Sarma *et al.* 2010). This biochemical features also support the importance of this ncRNAs in XCI.

TSIX *Tsix*, which is reported to be an anti-sense transcript to *Xist*, has been known to *Xist* repressor (Lee *et al.* 1999). Its candidate coding region has been suggested by one report which was observing that deleting 65 Kb of *Xist* 3' region in XO individual resulted in inactivation of the X-chromosome (Clerc and Avner 1998). This means the deleted region would be essential compartment for controlling the inactivation of X-chromosome by preventing *Xist*. The transcript which is expressed from *Xist* antisense strand is transcribed on active X-chromosome preferentially and diminished from the future inactive X-chromosome in differentiating mouse ES cells. Some studies revealed that this gene is related to the imprinted and random XCI in mouse (Lee and Lu 1999; Sado *et al.* 2001; Lee 2005). It was observed that the embryo obtaining mutated maternal *Tsix* undergone death because of upregulation of *Xist* allele in *Tsix*-mutated allele and follow by failure of forming extraembryonic development. In mouse ES cells, *Tsix*-mutated

X-chromosome showed preferential inactivation in mouse model. And also, *Tsix*-null mouse ES cells showed drastic expression of *Xist*. These result supports that *Tsix* has critical roles involved with the accurate regulation of *Xist* expression for choosing the future inactive X-chromosome in mouse XCI. Several studies suggested that the repression of *Xist* through *Tsix*. One report suggest that *Tsix* transcript prevent the loading of the *RepA* transcript connected with PRC2 protein onto the *Xist* promoter region and this would lead to the repression of inactivation of *Xist* expression (Zhao *et al.* 2008). Other reports showed that *Tsix* induced the methylated promoter region of *Xist* by recruiting DNMT3a (Navarro *et al.* 2006) or modifying the chromatin status in 5' region of *Xist* to repressed its expression (Sado *et al.* 2005; Sun *et al.* 2006). However, compared to the mice, which are required the accurate regulation of monoallelic *Xist* expression in developmental phase, other species like human and rabbit showed less coordinated monoallelic expression of *XIST* in their embryonic development because both species didn't showed tightly regulated monoallelic *XIST* expression (Okamoto *et al.* 2011). And although the counterparts of *Tsix* was reported in human (Migeon *et al.* 2001), the conserved functions and sequences are unclear in other eutherians (Chureau *et al.* 2002; Migeon *et al.* 2002; Chang and Brown 2010; Escamilla-Del-Arenal *et al.* 2011). Therefore, conserved function and orthologs of *Tsix* in other eutherians are needed to be examined further.

JPX The gene, which is also called *Enox*, is discovered at the about 10 Kb upstream from 5' region of *Xist* with opposite transcription strand (Johnston *et al.* 2002). This gene is considered to be originated from the ancestral protein coding gene, *USP1*, and the first exon of the *USP1* showed sequence homology with promoter region of human *JPX* gene. Most compartment of the ncRNA is mobile element and the sequences of the orthologs are variable among the species. The results of sequence comparison indicate that this gene is originated through species-specific manner like *XIST* orthologs (Chureau *et al.* 2002). Recent two studies clearly demonstrated that the gene is positive regulator for *Xist* expression in mice. Expression levels of the gene were increased in differentiating male and female mouse ES cells and this means that the gene is XCI-escaping gene and its expression is not affected by the *Xist*. Deletion of the gene was not influenced in the male ES cells differentiation but, in female, it led to cell death and down regulated expression of *Xist*. Exogenous *Jpx* expression in heterozygotic female ES cells showed rescued XCI rate and normal growth of the differentiating ES cells (Tian *et al.* 2010). Additional study revealed that the gene works with dose-dependent manner for inducing XCI and it is suspected that the transcript would compete with CTCF, which is *Xist* repressor by binding to the promoter region. This suggestion is originated by up- or down-regulation of *Xist* gene through the overexpression of CTCF or *Jpx*, respectively. Consistently, it is confirmed that CTCF could bind *Jpx* transcript (Sun *et al.* 2013). Therefore, *Jpx* is a *Xist*-activator and its functions are considered to be associated with CTCF closely. However, as

only a few studies were performed to confirm the roles of *Jpx* in mouse and the sequences of the orthologs are less conserved, it is required that the functions of the gene are examined in various species.

FTX This ncRNA is posited Five prime To *Xist* and considered to be evolved from ancestral gene called *WAVE4* (Elisaphenko *et al.* 2008). The five exons from 5' region of the gene showed conservation between human and mouse orthologs but the alternative spliced variants from the ortholog of each species showed differential construct (Chureau *et al.* 2002). As described previously, the ncRNA is also generally less conserved compared to the protein coding genes in XIC. However, interestingly, microRNAs, miR374a and miR545, are present within the *FTX* orthologs commonly among the most eutherian species. The function of the gene is firstly suggested in 2011 by analyzing the changes on *Ftx*-null mouse ES cells (Chureau *et al.* 2011). Deletion of the gene affected to its surrounding gene expression. The genes transcribed with same direction to *Ftx*, like *Xist*, showed down regulated but the others with the opposite transcribing direction was not affected following the deletion of *Ftx*. And it induced elevation of DNA methylation and reduction of H3K4ME2, which is active chromatin marker, in *Xist* promoter region. However, global histone modification for active or repressed chromatin was not affected by the *Ftx* removing, therefore, it is concluded that *Ftx* is positive regulator for *Xist* expression by regulating the chromatin in close

distance. Recent study observed that deletion of *Ftx* didn't affect to the imprinted XCI in preimplantation embryos (Soma *et al.* 2014). However, its roles in late stage of embryo development are not confirmed and only a few studies were suggested the possible roles of the gene, exact roles of this ncRNAs in imprinted or random XCI is required to be confirmed.

RLIM *RLIM*, also called *RNF12*, is E3-ubiquitin ligase regulating LIM-homeodomain containing factors like LDB1. This gene is highly conserved (Rastan and Robertson 1985) and not included in classical XIC suggested in its first identification (Rastan 1983). Recent studies have been suggested strongly that this gene is one of the important regulators for inducing XCI as a *Xist* activator (Jonkers *et al.* 2009; Shin *et al.* 2010; Barakat *et al.* 2011; Barakat *et al.* 2014). Transduction of exogenous BAC clones harboring this gene induced ectopic expression of *Xist* in male ES cells during their differentiation (Jonkers *et al.* 2009). This screening process resulted in identification of *Rlim* as a positive regulator for *Xist* expression. Function of the gene is conserved between mouse and human, and its heterozygote ES cells showed reduced XCI during their differentiation. These results demonstrate that this X-linked gene regulates the *Xist* expression depend on its dose. Subsequent study revealed that this gene directly regulate *Xist*, rather than *Tsix*, and XCI as a *trans* activator (Barakat *et al.* 2011). The study confirmed that *Rlim* bound to *Xist* promoter region. And also XCI in differentiating ES cells with

heterozygotes *Rlim* allele was rescued by the exogenous genes dose-dependently. This result exhibited the importance of the factor for inducing XCI by *trans* acting. The factor has been also considered to be an intermediate for molecular circuit of XCI because pluripotent factors, *Oct4*, *Sox2*, and *Nanog*, which are another regulatory factor for regulating *Xist* expression (Navarro *et al.* 2008), also target this X-linked *trans* activator to repress its expression (Navarro *et al.* 2011). Other reports confirmed the relation between *Rex1* and *Rlim* by observing *Rlim*-mediated *Rex1*-degradation (Gontan *et al.* 2012). Recent one study suggested that *RLIM* inactivated X-chromosome independent to physical interaction by pairing of two X-chromosome using heterokaryon ES cell model. These results obtained by using mouse ES cell model showed the critical function of the gene in random XCI. However, other two reports showed differential results to those in ES cell models by examining the function of *RLIM* using embryo models (Shin *et al.* 2010; Shin *et al.* 2014). Shin and his colleagues observed that *Rlim*-null preimplantation embryos failed to induce imprinted XCI and the fetus failed to form extraembryonic tissues (Shin *et al.* 2010). And also the ES cells without *Rlim* succeeded in initiation of XCI. Add to the results, post-implantation embryos lacking *Rlim* showed detectable *Xist* cloud and heterochromatin marker, H3K27me3, which are representing inactive X-chromosome marker (Shin *et al.* 2014). And also *Rlim* lacking embryos and ES cells which were assessed to tetraploid complementation showed embryogenic potential. These results demonstrate that *Rlim* is less related to random XCI and this is opposite to previous study (Jonkers *et al.* 2009; Barakat *et al.* 2011; Barakat

et al. 2014). Therefore, more studies are conducted to define the exact function of *Rlim* in XCI.

2.2 Pluripotent factor

Compared to the genes in XCI, pluripotent factors have been reported to be repressors for *Xist* expression and/or activators for *Tsix* expression. Considering that the close relation between maintenance of pluripotency in ES cells and XCI initiation during their differentiation, it is easy to expect that pluripotent factors are related to XCI closely. After the first identification of pluripotent factors, *OCT4*, which is gate-keeper for maintaining pluripotency in ES cells and ICM, (Nichols *et al.* 1998; Niwa *et al.* 2000), various factors were reported to regulate pluripotency and reprogramming. Most representatively, deriving the pluripotent stem cell artificially by reprogramming the differentiated cells through transduction of exogenous pluripotent factors (Takahashi and Yamanaka 2006) demonstrates the genes are closely related to the specific cell status showing pluripotency like ICM and ES cells. And also, reprogramming somatic cells to pluripotent cells coincided the reactivation of inactive X-chromosome (Ohhata and Wutz 2013). These tight linkages between *Xist* expression/XCI and pluripotency suggested the possibilities that pluripotent factors would act as a XCI-regulators. However, examining clear relations between XCI and pluripotent genes is a difficult procedure because the regulatory circuit of pluripotent marker is complex and each factor manages

various downstream targets. And also, as XCI and depletion of pluripotent marker is observed almost simultaneously during differentiation of ES cells, it is hard to distinguish obviously that the factors have genuine actions on XCI or both are independent outcomes resulted from differentiation. Nevertheless, several studies have examined the relation between this complex network after the first suggestion that represent pluripotent genes, *Oct4*, *Sox2*, and *Nanog*, regulate *Xist* expression in mice (Navarro *et al.* 2008). After the reports, subsequence studies were confirmed that the factors related to pluripotency were associated with XCI inducer, *Xist* or its negative regulator, *Tsix*. In this section, the roles of pluripotent factors in XCI by regulating *Xist* or *Tsix* expressions will be described.

OCT4, SOX2, and NANOG The functions of these three factors which is known to maintain pluripotency in ES cells (Nichols *et al.* 1998; Chambers *et al.* 2003; Masui *et al.* 2007) in *Xist* expression were examined in 2008 (Navarro *et al.* 2008). Binding of the factors on the first intronic region of *Xist* was observed in undifferentiated ES cells but they were detached when the cells are differentiated. Interestingly, *Tsix* and *Xist* showed mutually opposite expression patterns during differentiation, but the timing for expression level changes of *Xist* were more sensitive rather than that of *Tsix*. Therefore, this three pluripotent factor have been suggested as negative regulators for *Xist* expression. Another study revealed that *OCT4* and *SOX2* also regulate *Tsix* expression by recruiting and forming complex

with Ctf and Yy1 (Donohoe *et al.* 2009) which are *trans* factor regulating XCI (Chao *et al.* 2002; Donohoe *et al.* 2007). Consensus sequences of Oct4 and Sox2 were observed in close to the Ctf and Yy1 binding site in *Tsix* promoter region and *Xite* which is an enhancer for *Tsix* expression (Lee 2005), and their binding was proved *in vitro* and *in vivo*. Knockdown of *Oct4* and *Ctf* showed drastic effect on the pairing of X-chromosome which is observed prior to one of the X-chromosome inactivation, and aberrant expression of *Xist* in female ES cells compared to the downregulation of *Sox2* and *Yy1*. There results showed *Oct4* regulate *Tsix* expression critically. However, because of low binding levels of Oct4 and Sox2 to *Xite* which is a weak enhancer for *Tsix* expression compared to the *DXPas34* (Ogawa and Lee 2003; Vigneau *et al.* 2006), the functions of these pluripotent genes in *Tsix* expression were considered to be mild and it is possible that another factors were associated to the *Tsix* additionally or more importantly. And also, other studies suggested that deleting the first intron of *Xist* which was reported to be a target region of pluripotent factors, didn't affect to the inactivation and reactivation of X-chromosome (Barakat *et al.* 2011; Minkovsky *et al.* 2013). So it is controversial whether these represent pluripotent markers are powerful regulators for XCI repression. Another function of these factors in XCI is down regulation of *Rlim* which has been reported to a *trans* activator for inducing XCI (Navarro *et al.* 2011). Although there are uncertainties about how the factors affect the *Xist* and *Tsix* expressions, previous studies in ES cells show that these factors are XCI-repressors.

REX1 and PRDM14 Although one report addressed that *Oct4* is top regulator for XCI (Donohoe *et al.* 2009), several properties of *Tsix* enhancers and mild effect of *Oct4* repression in *Tsix* raised the possibility that other factors would be associated with *Tsix* expression additionally. Therefore, other three pluripotent factors, *Myc*, *Klf4*, and *Rex1* were examined to confirm whether they have relation with *Tsix* expression (Navarro *et al.* 2010). These three factors bound to the each end of *DXPas34* (*Rex1*) or between *DXPas34* and *Tsix* promoter (*Myc* and *Klf4*). Among them, displacement of the *Rex1* from *Tsix* regulatory sites in trophoblast cells showed reduced *Tsix* expression and repression. These result demonstrated that *Rex1* have important roles in *Tsix* expression. And also, *Rex1* binding to *Xist* promoter was also observed when the ligase function of the *Rlim* was prevented. And *Rex1* overexpression induced reduced XCI initiation in differentiating female ES cells but its repression in male ES cells resulted in forming ectopic *Xist* cloud. This means that *Rex1* is negative regulator for XCI by repressing *Xist* expression (Gontan *et al.* 2012). Therefore the factor would have dual functions, repressing or activating *Xist* or *Tsix*, respectively.

Prdm14, which is one of the subfamily genes of tandem zinc fingers and RB domain containing transcription factor, was reported to be an important regulator for XCI recently (Ma *et al.* 2011; Payer *et al.* 2013). This gene expressed in pluripotent cells like compacted morula, ICM, early epiblast cells, and primordial germ cells (Kurimoto *et al.* 2008; Yamaji *et al.* 2008; Yamaji *et al.* 2013) and interferences of the transcript resulted in differentiation of ES cells into

extraembryonic endodermal cells (Ma *et al.* 2011). Elevation of *Xist* expression was observed when *Prdm14* was downregulated in ES cells (Ma *et al.* 2011) but other report obtained the result that *Xist* expression in *Prdm14*-null mouse ES cells wasn't changed (Payer *et al.* 2013). Nevertheless the differential results within two studies, it is considered that this pluripotent factor would repress XCI. As other pluripotent markers, Oct4, Sox2, and Nanog (Navarro *et al.* 2008), *Xist* intron 1, rather than its promoter region, was target site of *Prdm14* and the region was close to the Nanog-binding site (Ma *et al.* 2011). Add to the binding site in *Xist* locus, promoter region of *Rlim* was also suggested to the target region for *Prdm14* by coordination of PRC2 and its subunit, SUZ12 (Payer *et al.* 2013). Indeed, ES cells derived from *Prdm14*-null mouse showed increased expression levels of *Rlim* compared to the wild-type female ES cells (Payer *et al.* 2013), and these results demonstrate that *Prdm14* has direct or indirect roles for repressing *Xist* expression via binding to *Xist* intron or silencing *Xist* activator, *Rlim* (Jonkers *et al.* 2009; Shin *et al.* 2010; Barakat *et al.* 2011). However, *Tsix*, the opposite regulator of *Xist*, was not considered to the target for *Prdm14* even though the gene support the reactivation of inactive X-chromosome in reprogramming epiblast stem cells (EpiSCs) to ES cells and establishing induced pluripotent stem cells (iPSCs) (Gillich *et al.* 2012; Payer *et al.* 2013) because deletion of the gene didn't affect to the *Tsix* expression and *Prdm14* was not bound to *Tsix* locus. Oppositely, *Tsix* would have effect on the localization of *Prdm14* to *Xist* intron 1 (Payer *et al.* 2013). These results showed that *Prdm14* act as a negative regulator for XCI by relating

expressions of *Xist* and its activators rather than that of *Tsix*.

When considering the expression of *Prdm14* and *Rex1* in developing embryos, their expression may have more relation to XCI compared to the other pluripotent factors like *Oct4*, *Sox2*, and *Nanog* as *Prdm14* and *Rex1* expression is diminished but others still strongly expressed in the epiblast. This means these two pluripotent factors of which expressions are reduced rapidly after differentiation of ICM into epiblast and hypoblast would be related powerfully to the XCI rather than *Oct4*, *Sox2*, and *Nanog*. However, as described above, it is hard to confirm the accurate association between pluripotent markers and XCI, therefore, more studies regarding the molecular networking for activating or inactivating XCI should be carried out.

As introduced above, numerous factors are involved with XCI. These factors form the delicate and tight networking for XCI by regulating XCI inducers or repressors like *Xist* and *Tsix* with direct or indirect manners (Figure II.3). XCI in *in vitro* (ES cells) or *in vivo* (developing embryos) models showed that the process is accomplished within less than five or three days, respectively and the complex and dynamic changes of molecular networking also observed in short period. However, unfortunately, the studies regarding the molecular regulatory were performed in mouse model mainly despite the differential strategies for XCI among the eutherians. Therefore the functions of each factor in other species are also

examined additionally even if pluripotent factors, which are one cluster of *trans* regulator for XCI, are generally conserved among the vertebrates.

3. Applications of XCI in embryo development and pluripotent stem cells

Although various studies have confirmed and examined the XCI using embryos and ES cells, several studies applied the mechanism for evaluating normal development of embryos, especially clones, and pluripotent status of stem cells. This application is due to that this epigenetic process in embryos is regulated very accurately and abnormal reprogramming in cloned embryos induces unregulated genes network by skewed XCI (Inoue *et al.* 2010). And also, the pluripotent stem cells like ES cells or EpiSCs reflect the status of XCI of each origin (Tesar *et al.* 2007), recent studies trying to establish ES cells in non-mouse models have applied the XCI status in obtained pluripotent stem cell to evaluate the origin and pluripotent status of them (Nichols and Smith 2009). The importance of the process has been focused recently in examining normality during developmental process and the status of pluripotency in ES-like cells, therefore, the studies regarding XCI in preimplantation embryos and pluripotent stem cells will be described in here.

3.1. Aberrant incidence of XCI in developmental competence

Even though numerous studies were carried out to confirm the XCI process and

the functions of its regulators using embryos, XCI status in *in vitro* embryos were examined in various species and applied as an indicator for the properly achieved reprogramming. After the first production of cloned animal using somatic cell nuclear transfer (SCNT) (Wilmut *et al.* 1997), the techniques has been applied to various species successfully for last 20 years. However, the developmental competence and birth rate of the cloned embryos are still low despite effort for improving the techniques. Various abnormalities in cloned embryos were observed during developmental period like large offspring or placental syndrome and after their birth with various pathologies (Young *et al.* 1998; Wilmut *et al.* 2002). Although the reasons of this recessive development would be variable, incomplete and unstable reprogramming of the donor cells has been suggested to be one of the most critical reasons for the abnormal developments in cloned animals. During the reprogramming process, hypomethylation of *Xist* promoter region in active X-chromosome showed inappropriate expressions of the *Xist*, and resulted in silenced X-borne genes in embryos (Panning and Jaenisch 1996). This means abnormal reprogramming of *XIST* or X-linked genes result in unregulated development. Thus, several studies have been observed the abnormal regulations of the XCI and expression patterns of the process regulators, like *XIST*, in cloned embryos during their developmental period.

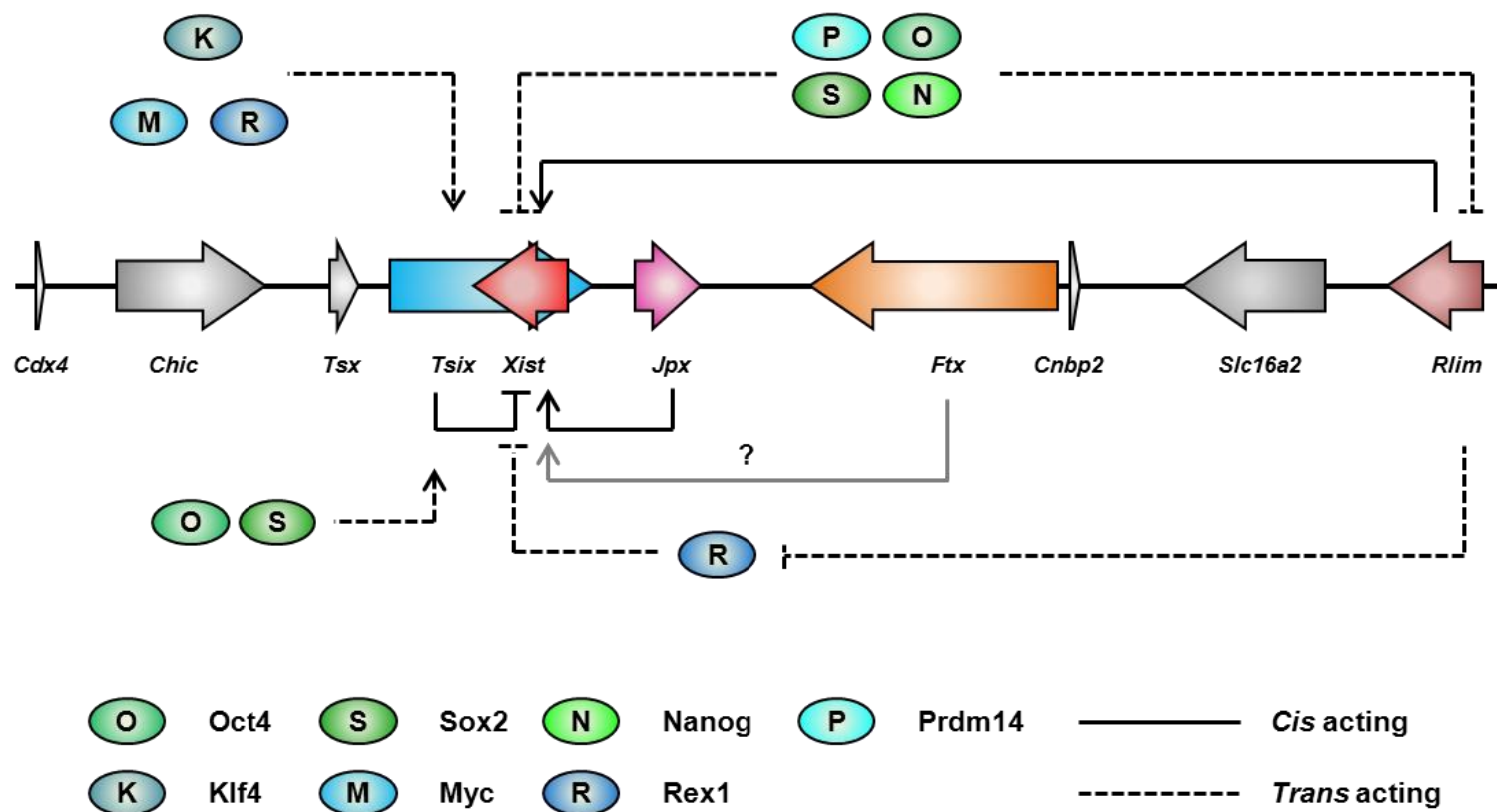


Figure II.3. Molecular networking of XCI-regulators, XIC-linked genes and pluripotent factors, in mouse XCI. *Cis* (solid lines) and *trans* (dashed lines) actions of XCI-regulators are depicted. Pluripotent factors generally act as a XCI repressors by down regulating *Xist* (*Oct4*, *Sox2*, *Nanog*, *Rex1*, and *Prdm14*) or up regulating *Tsix* (*Oct4*, *Sox2*, *Klf4*, *Myc*, and *Rex1*) expressions. XIC-linked genes are generally *cis* regulators for *XIST* expression. *Rlim* works as *cis* activator for *Xist* expression and also, *trans* repressor targeting *Rex1* by ligating ubiquitin to the protein. The positive functions of *Ftx* on *Xist* expression is still unclear.

Although cloned mouse showed imprinted XCI in their placental tissues following the donor cells and random XCI in their embryonic parts (Eggan *et al.* 2000), the number of live births was very small and some cell populations in cloned embryos showed heterogenic XCI patterns (Nolen *et al.* 2005). Indeed, even though cloned offspring showed normal XCI patterns, large number of deceased fetus couldn't sustain normal XCI (Xue *et al.* 2002; Nolen *et al.* 2005). Recent report showed that abnormal XCI by ectopic *Xist* expression was implicated in recessive development and birth of offspring in mouse cloning (Inoue *et al.* 2010). The study observed that expression levels of *Xist* was elevated in cloned mouse blastocyst and outstanding depression of genes linked to X-chromosome rather than autosomes. And also, one and two ectopic *Xist* clouds were observed in male and female blastocyst, respectively. This means cloned embryo undergone skewed XCI in active X-chromosome. This abnormal XCI and disrupted expressions of X-linked genes in cloned blastocyst was recovered by deleting one *Xist* allele and interestingly, this resulted in 10-times increase of birth rate. Subsequent study supported the RNAi-mediated *Xist* modulation in cloned embryos also recovered the defected cloned mice creation (Matoba *et al.* 2011). These results clearly demonstrate the skewed XCI and ectopic *Xist* expression originated by incomplete reprogramming impedes the normal development of cloned embryos. This abnormal XCI in cloned embryos and their full-term development has been also observed in other species.

Placenta from deceased cloned bovine fetus showed non-imprinted and hypomethylated patterns on their *XIST* promoter region (Xue *et al.* 2002; Su *et al.*

2011), and an aberrant expression patterns of the gene also elevated compared to the normal births (Dindot *et al.* 2004a; Su *et al.* 2011). Similarly, the aberrant expressions of X-linked genes were observed in the neonates cloned piglets (Jiang *et al.* 2008). These result showed the irregular regeneration of X-chromosome, as a resultant of incomplete reprogramming during SCNT, would induce lethality during full-term development of cloned embryos as in mouse. The abnormalities observed in deceased or live birthed cloned animal like hypomethylated *XIST* promoter and unregulated expressions of *Xist* and X-linked genes are considered to be evolved from earlier stage of embryo. Therefore, several studies examined the differences on the expression patterns of X-linked genes including *XIST* following the embryo origins (Wrenzycki *et al.* 2002; Zhou *et al.* 2008; Park *et al.* 2012). Although there are some differences on the results, generally common features were observed. Representatively, abnormal expression of *XIST* was most typical patterns of cloned preimplantation embryos. Even though the accurate processes of XCI and cytological expression patterns of *XIST* have not defined in the domestic animals, it is a reliable expectation that aberrant expressions of *XIST* in cloned embryos by unsuccessful reprogramming during genome activation would make problems on orchestrating the X-linked genes in developing process. And also, this would lead to the failure of compensating dosage of X-linked genes between male and females and might induce epigenetic disorders in clones. Therefore, the expression patterns of *XIST* and X-linked genes, and epigenetic status of XCI in developing embryos could be a good indicator for evaluating normality of

reprogramming process. However, even though the differences of XCI in the cloned and wild type individuals were observed, it is required that confirming the exact process of the XCI in each species to set the standards for distinguish normal or abnormal XCI in embryos.

3.2. Indicator for the status of pluripotency

As in studies regarding embryos and their full-term development, patterns of XCI have been used for one of the standard distinguishing the pluripotent status of stem cells derived from embryos with differential developmental stage. The stem cells from differential origins reflect their source cells, for example ICM and epiblast (Rossant 2008), therefore it is possible that the patterns of XCI in each stage of embryos could be an indicator for pluripotency of their derivate oppositely. This approach has been highlighted after the establishment EpiSCs in mouse (Brons *et al.* 2007; Tesar *et al.* 2007) although the ES cells have been used for the most popular models for XCI studies in mouse because of the epigenetic similarities between origins and derivate, These two types of embryo-derived stem cells showed differential properties each other (Nichols and Smith 2009). Briefly, mouse ES cells, also called 'naïve' pluripotent stem cells, are originated from early epiblast (ICM) and form compact colonies with dome-shape. And also, their self-renewal and maintain are sensitive to leukemia-inhibition factor (LIF) and mainly respond to LIF-STAT3 pathway. However, EpiSCs are originated from late stage of

epiblast and their colonies showed flatten shape with loose density. The cells apply FGF-ERK pathway, which induce the differentiation of ES cells, for their proliferation. Differences on chimera and teratoma formation ability were also observed between these two cell lines. Add to the differential epigenetic properties between two types of pluripotent cells, inactivation status of X-chromosome is also different. When considering the dynamics of X-chromosome status during embryonic stage, the difference is natural because the double active Xs is observed at the early epiblastocyst stage by reactivation of imprinted Xp and followed by being inactivated randomly in late epiblast (Heard 2004; Okamoto *et al.* 2004; Ohhata and Wutz 2013). Therefore, XCI status in pluripotent stem cells has been applied as one of the markers for distinguishing naïve and primed pluripotent stem cells (Payer *et al.* 2011). Interestingly, mouse ES cells and iPSCs reprogrammed from EpiSCs and somatic cells showed reactivated inactive X-chromosome (Maherali *et al.* 2007; Guo *et al.* 2009), and these results also support that this epigenetic mark is proper to judge the status of pluripotent cells and developmental stage of counterpart origin.

Pluripotent stem cells from blastocyst or reprogrammed somatic cells in non-rodent species showed similar features to the mouse EpiSCs, like flatten colonies, required cytokine for maintaining their self-renewal, and reduced abilities on chimera and teratoma formations (Nichols and Smith 2009; Nowak-Imialek and Niemann 2012; Koh and Piedrahita 2014). Some studies suggested that ES-like cells and iPSCs in human and ungulates would be primed status by showing one

inactive X-chromosome (Tchieu *et al.* 2010; Park *et al.* 2013b) although human ES cells showed heterogenic cell population with differential XCI patterns (Silva *et al.* 2008; Tanasijevic *et al.* 2009). And also, recent studies succeed establishment of naïve-like pluripotent stem cells having reprogrammed two active X-chromosomes in non-mouse models by modulating cell signaling or exogenous factors (Hanna *et al.* 2010; Fujishiro *et al.* 2013; Gafni *et al.* 2013). These results support that the XCI status in pluripotent stem cell is outstanding indicators for distinguishing their pluripotent status.

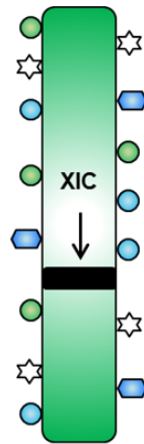
4. Techniques for XCI analysis

XCI is complex epigenetic process with various activators and repressors mentioned above. And also, the status of XCI evaluation is hard to be determined just by examining expression of *XIST* because the completion of the XCI is required several epigenetic changes on the future inactive X-chromosome gradually (Escamilla-Del-Arenal *et al.* 2011). Therefore, the inactivation mechanisms of X-chromosome will be described first (Figure II.4) and various methods for examining XCI status focusing represent marks of inactive X-chromosome will be introduced.

4.1. Epigenetic changes of X-chromosome in its inactivation

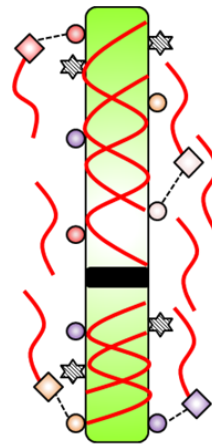
***XIST* RNA coating** The procedure of XCI is initiated with accumulation of *Xist* transcript on one of the X-chromosomes. Among the repeat region in *Xist* transcript, repeat C has reported to manage localization of *Xist* to X-chromosome by interacting with RGC domains in hnRNP U, SP120/SAF-A proteins in mice (Hasegawa *et al.* 2010; Sarma *et al.* 2010). As following the elevation of localized *Xist* RNA to X-chromosome, the chromosome is coated with the transcript and forms secondary structure with encompassing repeat sequences like LINE in future inactive X-chromosome. During this process, the genes subjected to be inactive localize into the compartment of silent *Xist* RNA and XCI-escaping genes were moved to edge or outside from the *Xist* RNA domain (Chaumeil *et al.* 2006; Eymery *et al.* 2009). This *Xist* coating in X-chromosome shows loss of RNA polymerase II, Tbp, and Cot1 from the chromosome and removal of euchromatin histone modifications (Chaumeil *et al.* 2002; Okamoto *et al.* 2004; Chaumeil *et al.* 2006).

Early chromatin modifications After *Xist* coating on the X-chromosome, chromatin modification occurs dynamically and gene expression in X-chromosome is rapidly silenced in differentiating mouse ES cells (Escamilla-Del-Arenal *et al.* 2011). The first event in early phase of chromatin modification (first 2-3 days after differentiation of mouse ES Cells) is erasing euchromatin histone marks.



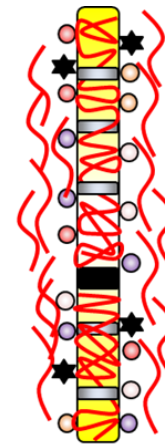
Undifferentiated ES cells

- Euchromatin
- Active gene expression



Differentiating ES cells (Day 2-4)

- Heterochromatin forming
- Gene silencing initiation
- Repressive chromatin modification



Differentiating ES cells (Day 4-8)

- Condensed chromatin
- Extension of heterochromatin marks
- Stable Inactivation (Irreversible)

	H3K4ME2/3		H3K36 methylation		H3/H4 acetylation		<i>Xist</i>
	H3K27ME3		H3K9ME2		H4K20ME1		H2AK119ub
	PRC2		Histone methyl-transferase		Pr-Set1		PRC1
	Active gene		Silenced gene (Non-methyl-promoter)		Silenced gene (Methyl-promoter)		MacroH2A

Figure II.4. Chromatin modification and inactivation process of X-chromosome in differentiating mouse ES cells. The epigenetic changes of future inactive X-chromosome in mouse ES cells are represented. Euchromatin marks like methylated H3K4, H3K36 and acetylated H3 and H4 are observed in the X-chromosome of undifferentiated mouse ES cells. However, the marks are replaced to heterochromatin marks, methylated H3K27, H3K9, and H4K20 and ubiquitinated H2AK11, after initiating differentiation. *Xist* coating and condensation of chromosome is also observed in differentiating ES cells. Heterochromatin marks are extended and irreversible inactive X-chromosome is formed at the four to eight days later from the differentiation.

H3K4ME2/3, which are commonly located promoter region and induce active expression of genes (Heintzman *et al.* 2007), are removed following *Xist* coating (Heard *et al.* 2001; Chaumeil *et al.* 2002). And acetylation on the H3/H4 was also removed (Chaumeil *et al.* 2002). After the removal of these euchromatin marks, heterochromatic marks are recruited on the inactive X-chromosome. Plentiful H3K27, H3K9ME2, H4K20ME1, and H2AK119ub are observed in the inactive X-chromosomes (Escamilla-Del-Arenal *et al.* 2011). H3K27ME3 is abundant in inactive X-chromosome and generally located on the promoter region of X-linked genes which are inactivated (Marks *et al.* 2009). This post-translational modification is formed by PRC2, which consist of three constituents, histone methyltransferase (Ezh2), zinc-finger containing protein (Suz12), and WD40 repeat protein (Eed) (Margueron and Reinberg 2011). Rapid recruitment of the complex was observed in *Xist*-coated X-chromosome of differentiated ES cells and simultaneously, accumulation of H3K27ME3 also occurred (Plath *et al.* 2003; Silva *et al.* 2003). Although the exact mechanism how to PRC2 is recruited to inactive X-chromosome to pass the methyl-group to the histone protein is unclear, it is thought that *Rep A*, which is 1 - 1.5 Kb of RNA transcribed from *Xist* repeat A domain (Jeon *et al.* 2012), would interact directly with PRC2 complex and resulted in H3K27ME3 enrichment (Escamilla-Del-Arenal *et al.* 2011) because the RNA could bind to Ezh2 and Suz12 in ES cells (Kaneko *et al.* 2010; Kanhere *et al.* 2010). Other heterochromatin marks, H3K9ME2, H4K20ME1, and H2AK119ub are also observed in *Xist*-coated X-chromosomes and they would be interacted by histone

methyl transferase, Pr-Set1, and PRC1, respectively (Escamilla-Del-Arenal *et al.* 2011).

Late chromatin modifications The modification is processed to establish the stable and irreversible inactive X-chromosome (after 4 – 8 days from mouse ES cells differentiation) (Marks *et al.* 2009). The modifications in early phase of inactive chromatin are maintained or extended in this step (Escamilla-Del-Arenal *et al.* 2011). Represent differences in this stage are DNA methylation on the promoter and CpG sites of X-linked genes and enrichment macroH2A protein in inactive X-chromosome (Norris *et al.* 1991; Tribioli *et al.* 1992; Costanzi and Pehrson 1998; Chadwick *et al.* 2001). Although the global DNA methylation levels are higher in active X-chromosome (Weber *et al.* 2007) but methylated CpG sites and promoter are preferentially observed in the X-linked genes on inactive X-chromosomes (Norris *et al.* 1991; Tribioli *et al.* 1992). This epigenetic mark on promoter region is considered to be important process for sustaining stable inactive status and silencing transcription of X-linked genes. Plentiful macroH2A, which is variant of H2A having H2A-like domain with non-histone part (Pehrson and Fried 1992), was observed in inactive X-chromosome (Costanzi and Pehrson 1998; Chadwick *et al.* 2001). Considering its functions repressing RNA polymerase transcription (Doyen *et al.* 2006) and global distribution in inactive X-chromosome (Mietton *et al.* 2009), macroH2A is thought to be related with formatting inactivation chromatin structure

and resulting in silencing X-linked genes stably. Elevation of several factors like Ash2l, Atrx, and hnRNP U, compared to the earlier phase heterochromatin are also observed in stabilizing inactive X-chromosome (Baumann and De La Fuente 2009; Pullirsch *et al.* 2010) but their functions in stabilizing inactive chromatins are unclear.

4.2. Cytological analysis for evaluating XCI status

Cytological analysis of XCI are mainly performed by applying RNA fluorescent *in situ* hybridization (FISH) of ncRNAs like *XIST*, and observing distribution of histone modifications like H3K27ME3 which are increased after XCI (Figure II.4). The advance of the approach is confirming the chromatin changes directly. Although the *XIST* is a main factor for XCI, it is hard to determine XCI just observing its expression using PCR because active X-chromosome showed basal expression of *Xist* RNA in mouse ES cells and PCR couldn't distinguish biallelic or monoallelic expressions. Therefore, this cytological approach based on the RNA FISH has been used for XCI studies. Fluorescent-tagged probes which are antisense to target transcript and could hybridize with the RNA are used to detect ncRNA. Inactivation of X-chromosome is distinguished by the staining patterns of *XIST* RNA, cloud pattern and pin-point pattern. Cloud and pin-point staining patterns indicate accumulated *XIST* transcript in inactive X-chromosome and basal expression of the gene in active X-chromosome, respectively. If two X-

chromosomes showed polymorphism on *XIST* allele, like repeat region, it is possible to confirm the expression origin also. In fertilized embryos, which is hard to distinguish sex, DNA FISH targeting X or Y-chromosome is conducted concurrently with RNA FISH (van den Berg *et al.* 2009). Distribution of heterochromatin marks in inactive X-chromosome, mostly H3K27ME3, is also applied to confirm the XCI with *XIST* RNA FISH. In this case, strong signal of H3K27ME3 is observed in inactive X-chromosome but it is hard to conclude that cells showing intensive signal of the histone modification is XCI-completed cells because the signal could be observed also in other chromosomes. Therefore, co-localization of H3K27ME3 spot and *XIST* coating indicates inactive X-chromosome (Okamoto *et al.* 2011). The expression of X-linked gene is also used for the determining chromosomal silencing. It is possible that expression of X-linked genes even which are subjected to the XCI are detected on the *XIST*-coated X-chromosome. As the silencing of X-linked genes during XCI is achieved gradually, it is also required that examining expressing allele of the gene. Another merit of this analysis is discriminative examination of XCI in cells with different lineage. For example, XCI in ICM or trophoblast could be determined discriminatively by detecting each cell-specific marker together (Okamoto *et al.* 2011). However, as it is hard to examining expression of large number of X-linked genes, global silencing of each X-linked gene is hard to be confirmed using the method. And also, the RNA FISH is highly influenced to the probes of target genes, the differential results could be obtained in same type of cells (van den Berg *et al.*

2009; Okamoto *et al.* 2011).

4.3. Non-cytological analysis for evaluating XCI status

Although RNA FISH with observing chromatin modification marks is considered generally proper method for the analysis of XCI, non-cytological analyses were used in several species of which exact sequences of *XIST* has not been defined. Expression of *XIST* and other X-linked genes has been examined using conventional methods, like PCR, but these results are hard to know that which chromosome is origin of the RNA expression. However, despite the limits in this indirect method for analyzing XCI, recent one study analyzed the expression levels of X-linked genes between male and female bovine blastocyst, and examined the chromosome-wide dosage compensation indirectly (Bermejo-Alvarez *et al.* 2010). Although the experiment couldn't distinguish each expression pattern in their two cell compartments, ICM and trophoblast, this study could concluded compensation status of each X-linked gene in blastocyst. This approach make possible confirm the gradual spreading of silenced genes in X-chromosome during their developmental period. Add to the global transcription analysis, examining methylation patterns of *XIST* and X-linked genes are other methods for evaluating XCI by confirming gene silencing. Because thoroughly inactivated X-chromosome showed methylated promoter and CpG sites of X-linked genes generally (Norris *et al.* 1991) and those of *XIST* is demethylated in inactive X-chromosome (Hwang *et*

al. 2013a), this analysis also used for evaluating accomplished XCI. Therefore, these non-cytological analyses are used as alternative methods for analyzing XCI status instead of cytological approach in non-mouse species whose experimental supports and background are restricted, even if they reveal the status indirectly.

CHAPTER III

**Identification of the porcine *XIST* gene and its
differential CpG methylation status in male and
female pig cells**

1. Abstract

XIST, a long non-coding RNA, plays an important role in triggering X chromosome inactivation in eutherians, and is used extensively for qualifying stem cells and cloned embryos. However, a porcine *XIST* has not yet been thoroughly identified despite its biological importance in a wide variety of research fields. Here, full-length porcine *XIST* sequence assembled using known sequences (GenBank), RNA-Seq data (NCBI SRA), and PCR/sequencing was represented. The proposed porcine *XIST* gene model encodes a 25,215-bp transcript consisting of 7 exons, including two conserved and two porcine-specific repeat regions. Transcription covering the entire *XIST* region was observed specifically in female cells, but not in male cells. Eight transcription starting sites (TSSs) were discovered and evaluated CpG methylation patterns in the upstream (+2.0 kb) and downstream (-2.0 kb) regions. Sixty-seven CG di-nucleotides identified in the target region were considered to be candidate CpG sites, and were enriched in the following two regions: -284 to +53 bp (13 sites) and +285 to +1,727 bp (54 sites) from the selected TSS. Male 5' region of *XIST* (64.5 sites, 96.26 %) had a higher level of CpG methylation than female DNA (33.4 sites, 49.85 %). Taken together, the results revealed that the porcine *XIST* gene is expressed exclusively in female cells, which is influenced by the lower level of CpG methylation in the putative promoter region compared with male cells. The porcine *XIST* presented in this study

represents a useful tool for related research areas such as porcine embryology and stem cell biology.

Key words: X chromosome inactivation, XIST, CpG site, Repeat sequence, Gene mapping

2. Introduction

X-chromosome inactivation (XCI) occurs early in embryo development, and involves silencing one of the X chromosomes in mammalian female cells to compensate gene dosage (Lyon 1961). Different mammalian species use various strategies to induce XCI during early embryo development (Okamoto and Heard 2009). For example, it has been suggested that eutherians except rodents, like humans, employ random inactivation of the X chromosome (Brown and Robinson 1997; Okamoto and Heard 2009), whereas marsupials and rodents undergo imprinted inactivation of the paternal X chromosome during early embryo development (Sharman 1971; Huynh and Lee 2003). The X-chromosome inactivation-specific transcript, *XIST/Xist*, which is located at the X-inactivation center (Xic), has been suggested to trigger XCI in eutherian mammals by cis-coating one of the X chromosomes in female mammals (Penny *et al.* 1996; Wutz and Jaenisch 2000).

A gene model of *XIST/Xist*, which produces a long, non-coding RNA (ncRNA) (Brockdorff *et al.* 1992; Brown *et al.* 1992), showed that *XIST/Xist* consists of large first and last exons with several small exons between these large exons in both humans and mice. Identification of the *XIST* gene was a turning point in research on XCI in various eutherians. For example, the identification of *XIST* allowed for

the evaluation of XCI in human, murine, and bovine species during early embryo development (Kay *et al.* 1993; Brown and Robinson 1997; De La Fuente *et al.* 1999). Recently, the importance of *XIST* RNA for early embryo development and embryonic stem cell research has been highlighted. For example, the identification of abnormal expression of *Xist* and X-chromosome-linked genes in a cloned mouse was reported (Inoue *et al.* 2010), and RNA interference (RNAi)-mediated downregulation of *Xist* during embryo development in cloned mice was shown to dramatically elevate birth rates (Matoba *et al.* 2011). These reports suggest an important role for *Xist* in early embryo development. It has also been shown that *Xist* expression is associated with the pluripotency markers Oct4, Sox2, and Nanog, and that lower expression of *Xist* is maintained in undifferentiated mouse embryonic stem cells (Navarro *et al.* 2008). *XIST* research has also focused on the status of pluripotency, specifically naïve and primed status, because these two different stem cell types exhibit different patterns of X chromosome inactivation (Nichols and Smith 2009). Therefore, *XIST* has been used as a marker for evaluating the status of pluripotent stem cells (Nichols and Smith 2009). Although the importance of *XIST* has been highlighted in various research fields, only partial *XIST* sequences have been reported in pig (Cepica *et al.* 2006). And also, putative gene model resulted from aligning between pig genome and bovine *XIST* was suggested recently, only small fraction compared to the suggested draft gene model was validated just using RT-PCR analysis (Bischoff *et al.* 2013). A recent paper that defined various genes and RNAs in the pig also identified partial sequences of

porcine *XIST*, namely, 84 bp of exon 1 and 126 bp of exon 4 (Groenen *et al.* 2012). Although these findings suggest some information about porcine *XIST* RNA, but a large proportion of *XIST* RNA sequence was still unknown. While the identified porcine *XIST* sequences could be used for research related to early embryo development and stem cells, identification of a full *XIST* sequence using gene prediction methods may be difficult (Groenen *et al.* 2012) because of the low sequence conservation of *XIST* genes among species (Hendrich *et al.* 1993; Pang *et al.* 2006).

The identification of *XIST/Xist* has also led to verification of its promoter region (Pillet *et al.* 1995; Hendrich *et al.* 1997) and analysis of the CpG dinucleotide methylation patterns in this region (Norris *et al.* 1994). The *XIST/Xist* promoter region has been suggested to regulate *XIST* expression by recruiting transcription factors (Sheardown *et al.* 1997b), and the methylation pattern of the promoter region of *Xist* is highly related to its expression (Norris *et al.* 1994). Analysis of methylation patterns has led to the suggestion that expression of *Xist* in the mouse is controlled by imprinting during preimplantation embryo development (Norris *et al.* 1994; McDonald *et al.* 1998). In addition to the identification of CpG sites as differentially methylated regions (DMRs) and factors in CpG sites in the *Xist* promoter region have been used to determine the normality of early embryo development (Nolen *et al.* 2005). Specifically, abnormal methylation patterns in the *Xist* promoter have been reported in cloned mouse embryos (Nolen *et al.* 2005). Nevertheless, little is currently known about the CpG sites of porcine *XIST*.

In this study, I aimed to identify the candidate porcine *XIST* ncRNA encoding regions using BLAST homology searches. I verified the findings using reverse transcription (RT)-PCR, enzyme mapping, sequencing, and RNA-Seq alignment on the whole region. Based on the identified *XIST* sequence, the methylation patterns of 67 CpG sites in the region ± 2 kb from the transcription start site (TSS) were analyzed to confirm the DMRs of porcine *XIST*. These results showed that the porcine *XIST* RNA is nearly 24 kb in length, and contains previously reported partial sequences of porcine *XIST*. The examined 67 CpG sites exhibited sex-dependent methylation patterns similar to those found in other species.

3. Materials and Methods

The pig experimental processes were performed strictly in accordance with the Guide for Care and Use of Research Animals in Teaching and Research. The experiment was approved by the Institutional Animal Care and Use Committee of Yonsei University, Wonju, Republic of Korea.

BLAST search

The entire pig genomic sequence (*Sus scrofa* genome, version 10.2) was compared to the exons of mouse *Xist* and human and bovine *XIST* (Accession Nos: NC_000086.7, NC_000023.10, and AC_000187.1, respectively) using BLAST. Regions on the pig X chromosome that gave an alignment score of over 300 and that matched at least two *Xist* or *XIST* sequences from other species were considered to be candidate encoding regions of the porcine *XIST* ncRNA.

Sample preparation, RNA extraction, and reverse transcription

The porcine embryonic fibroblasts (PEF, mixed breed) were obtained from the days post coitus (dpc) = 27 fetuses. Cultured male and female PEFs (passage=5)

were used in this study. PEF was cultured in Dulbecco's modified Eagle's medium (DMEM, Gibco Invitrogen, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum (FBS; collected and processed in Canada; Hyclone, Logan, UT), 2 mM glutamax (Gibco Invitrogen, Carlsbad, CA, USA), 0.1 mM β -mercaptoethanol (Gibco Invitrogen, Carlsbad, CA, USA), and 1x antibiotic/antimycotic (Gibco Invitrogen, Carlsbad, CA, USA) as previously described methods (Park *et al.* 2013a). Total RNA was extracted from cultured PEFs using TRIzol (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. Extracted RNA was diluted with sterile DEPC-treated water (pH=8.0) and treated with recombinant DNase I (Takara Bio, Otsu, Shiga, Japan). DNase-treated RNA samples (2.5 μ g) were converted to cDNA in 20 μ l final volume using Superscriptase III (Invitrogen Carlsbad, CA, USA) following the manufacturer's instrument.

PCR amplification

One-micro litter of synthesized cDNAs was subjected to PCR using the primers listed in Table III.1. PCR amplification was carried out using a 2x PCR master mix solution (iNtRON Bio Technology, Seongnam, Gyeonggi, Korea) containing 0.5 μ M of each primer set in 10 μ l reaction volume. PCR was performed according to the following conditions: 1 cycle of 95°C for 7 min; 35 cycles of 95°C for 45 sec, annealing at the temperature listed in Table III.1 for 30 sec, and 72°C for 1 min 30 sec; and 1 cycle of 72°C for 10 min. Amplification of *SRY* and *GAPDH* primer

pairs was performed according to the following conditions: 1 cycle of 95°C for 5 min; 30 cycles of 95°C for 30 sec, 60°C for 30 sec, and 72°C for 45 sec; and 1 cycle of 72°C for 7 min.

RACE-PCR of 5'- and 3'- region

All 5'- and 3'-region amplifications were performed using rapid amplification of cDNA ends (RACE) PCR with the 5'-RACE Core Set and 3'-RACE Core Set (Takara Bio, Otsu, Shiga, Japan), respectively. The process for synthesizing cDNA was carried out in accordance with the manufacturer's instructions and the amplification of synthesized cDNA samples were performed with modifications. Reverse-transcription and amplification was performed using the primer pairs listed in Table III.2. For the 5' RACE PCR, the first round of PCR was performed according to the following conditions: 1 cycle of 95°C for 7 min; 54°C for 5 min; 72°C for 40 min; 30 cycles of 95°C for 40 sec, 54°C for 1 min 30 sec, and 72°C for 2 min; and 1 cycle of 95°C for 40 sec, 54°C for 1 min 30sec, and 72°C for 15 min. The second round of PCR was carried out as follows: 1 cycle of 95°C for 7 min; 30 cycles of 95°C for 35 sec, 56°C for 30 sec, and 72°C for 1 min; and 1 cycle of 72°C for 10 min. In case of 3' RACE PCR, the first round of PCR was performed as follows: 1 cycle of 95°C for 5 min; 55°C for 5 min; 72°C for 40 min; and 20 cycles of 95°C for 40 sec, 55°C for 1 min, and 72°C for 2 min; and 1 cycle of 95°C for 40 sec, 55°C for 1 min, and 72°C for 15 min. Nested PCR was performed

according to the following conditions: 1 cycle of 95°C for 5 min followed by 35 cycles of 95°C for 30 sec, 60°C for 30 sec, and 72°C for 1 min 30 sec.

Enzyme cutting, cloning, and sequencing

Gel-extracted amplicons were purified using a commercial spin column (MEGAquick-spin™ Total Fragment DNA Purification Kit; iNtRON Bio Technology, Seongnam, Gyeonggi, Korea) for restriction enzyme mapping and cloning. Each purified PCR product was incubated with 2 U of the appropriate restriction enzyme listed in Table III.3 at 37°C for 2 hours. Amplicons were then cloned into the pGEM-T Easy vector (Promega, Madison, WI, USA) and transformed into *E. coli* (DH5α strain, Novagen, Madison, WI, USA). Plasmid samples were prepared using commercial spin columns (DNA-spin™ Plasmid DNA Purification Kit; iNtRON Bio Technology, Seongnam, Gyeonggi, Korea). Purified plasmid samples were sequenced with an ABI PRISM 3730 DNA Analyzer (Applied Biosystems, Foster, CA, USA) and compared with genomic scaffold sequences.

Genomic DNA extraction and bisulfite sequencing

Extraction of gDNA was performed using the G-spin Genomic DNA Extraction Kit for cell/tissue (iNtRON Bio Technology, Seongnam, Gyeonggi, Korea)

according to the manufacturer's instructions. Extracted gDNA was treated with bisulfite using the EZ DNA Methylation-Gold Kit (Zymo Research, Irvine, CA, USA) for methylation analysis of CpG dinucleotides according to the manufacturer's protocol. Bisulfite treated gDNA was amplified with primer pairs listed on the Table III.4 and a total of 67 CpG sites were analyzed following the condition. The first-round PCR was performed using 1 cycle of 95°C for 5 min and 35 cycles of 95°C for 30 sec, annealing at the temperature listed in Table III.4 for 30 sec, and 72°C for 2 min; and 1 cycle of 72°C for 10 min. Second-round PCR was performed as follows: 1 cycle of 95°C for 7 min; 40 cycles of 95°C for 35 sec, annealing temperature for 30 sec, and 72°C for 1 min; and 1 cycle of 72°C for 10 min. BS3 was amplified according to the following conditions: 1 cycle of 95°C for 5 min; 40 cycles of 95°C for 30 sec, 54°C for 30 sec, and 72°C for 2 min; and 1 cycle of 72°C for 10 min. Amplicons were cloned and sequenced as described above for enzyme cutting, cloning, and sequencing.

Illumina sequences, sequence alignment, and alignment visualization

Released Illumina sequences from one pair of female pig full-siblings (240 days old, breed: White Duroc X Erhualian F2) stored in the NCBI Sequence Read Archive (SRA) database (Accession Nos. SRX054582 - SRX054587) were used. The reads contained paired-end 90-bp sequences (200-bp insert) generated using the Illumina HiSeq 2000 platform for three different tissues—liver, abdominal fat,

and the longissimus dorsi muscle. Custom Perl scripts were used to enrich for higher quality (Q20) reads, which were subsequently aligned to *XIST* RNA sequences assembled manually using Bowtie2 (Langmead and Salzberg 2012). Aligned reads were sorted using SAMtools (Li *et al.* 2009) and visualized with Tablet (Milne *et al.* 2010).

Table III.1. The List of primer pairs sequence for reverse transcription PCR			
*Primer/Gene (Accession. No)	Sequence (5' → 3')	Tm (°C)	†Expected length (bp)
α	F : TACACTTTCCAGTTCCGCCA R : AGGTATCCACAGCCCCGA	64 °C	1427 bp
E1S3	F : TTTTCCCCCTCCTTTTCTGT R : TGACCCCGTTTTTCCACCTC	58 °C	799 bp
E1S1	F : AGGCTGGGGAGGAAAGATGG R : GGTGGTAATGCTGGAGGGGA	60 °C	953 bp
E1B1	F : CAGACAGTTAGTGGAGGATGGAA R : AAGGATGGGAACAAAGCAAGA	60 °C	899 bp
E1S2	F : AACAGGTGGCGGAAGAGGAA R : AAACAGGAGTGGGACAGGGG	64 °C	840 bp
E1S4	F : AGTGGGTTTTTCATTTTGGG R : GGGTCTTATCTGGTAGGCA	60 °C	1429 bp
E1N1	F : GGCATCCTACCATTTTTACCCCTCT R : AAAGTCCCTTGTATCCTCCCTTG	66 °C	1512 bp
E1N2	F : GTCTCACACTCAATAACCGCCT R : TACTCTCAGCAGCCGTCCTAAA	66 °C	1084 bp
E1A	F : AGAAGAAAGGGTGGGGGAAAAAC R : ACGGGAGGGGTCAGTAGAGCA	66 °C	1649 bp
E1B	F : ATTCCCGTTCCCTTACTCGTTTT R : TCCCTTGTATCTTCCTTGGTTGG	66 °C	1676 bp
E1C	F : ATTACCACAGAGGGGACAAGGG R : GCAAATCCAGTAGGACAGCAACA	66 °C	1426 bp
E1D	F : TACAAGGGACTGTTGATGGGCTT R : TGAAAATAGAGTAGGAGGGAGGGG	63 °C	1508 bp
E1E	F : CCCATCCCCATTCAACTTCC R : GCGACCTACACCACAGCCC	60 °C	1355 bp
E1F	F : TTCCCACCTAACTCCCTTTTCCT R : ACAGTCCCAAATGCCCTCCC	66 °C	1664 bp
E1G	F : GCCCTCTAACTTTTTACATTACCCC	66 °C	1706 bp

	R : GCACAAGAACCCAGACAAATACATC		
E1H	F : GCCTTGTCTTGGGACTGTTACTATG R : GTATGGGCTGCTGTTTGATGGA	64 °C	1455 bp
E2	F : CATCAAACAGCAGCCCATACTC R : ATCATTTCTAAGCCCTCACTTCAG	63 °C	1390 bp
E2EL **	F : GGAGCATCAACCAGCCCC R : GACACAGAAGCATACAAAGCACGA	63 °C	9538 bp (1868 bp)
ELB1	F : TTTCGTGCTTTGTATGCTTCTGTG R : GAGTAAGGTGTTGCTGGCTGATG	64 °C	1014 bp
ELS1	F : GAAGCACCAAGACCAAGGGAA R : CCACAGGCACAACAACGAGG	64 °C	1535 bp
ELB2	F : TTCCTTCATTTCTTTCCTCTTACC R : AGTCTTTCATTCATCAGGCATTT	60 °C	1697 bp
ELS2	F : GTGAGGCAGGCATTATCTCTACAA R : CCCATCTCGTCAATCAGGCA	60 °C	980 bp
ELB3	F : CTTGCTCCTTTGCCTGATTG R : TTTCACTCCATTTCTTCTTACTGTTG	58 °C	1071 bp
ELS3	F : TGCCAGAGAGTCAGAAAGCCA R : GAAATGAGGGGGAACAAAGGAA	58 °C	1571 bp
ELS4	F : CAGAGAGGGTGGGATAGGAGGAAG R : TGTTTCACATCAGTTCACAAGTCCA	58 °C	860 bp
β	F : CTCCTTCTTGAGGTGGGGGT R : GCTCCTGCTTGGTAATGGG	64 °C	860 bp
<i>SRY</i> (NM_214452.3)	F : CTGGGATGCAAGTGGAATAAT R : GGCTTTCTGTTCCCTGAGCAC	62 °C	250 bp
<i>GAPDH</i> ** (NM001206359)	F : GGTCGGAGTGAACGGATTTG R : GCCGTGGGTGGAATCATACT	62 °C	276 bp (174 bp)

*Primers were designed to result from BLAST analysis and the homologue region in *sus* scrofa 10.2 scaffold sequence was used (GenBank acc. no. NW_003612825.1) except *SRY* and *GAPDH*.

**The primer sets were designed to span intron

†The length in parenthesis represent the size of cDNA PCR amplicons

Table III.2. List of restriction enzymes used for each amplicon and the predicted sizes of the digested fragments.

Amplicon	Restriction enzyme	Predicted amplicon size (bp)	Digested fragment size (bp)*
E1S3	SacII	799 bp	144 + 655
E1S1	BstUI	953 bp	686 + 267
E1B1	EcoRI	899 bp	599 + 300
E1S2	EcoRI	840 bp	199 + 641
E1S4	DraI	1429 bp	208 + 849 + 372
E1N1	EcoRV	1512 bp	418 + (1094 + \approx 500)
E1N2	DraI	1084 bp	159 + (\approx 400) + (925+ \approx 200)
E1A	BamHI	1649 bp	989 + 660
E1B	ScaI	1676 bp	1189 + 487
E1C	ScaI	1426 bp	296 + 675 + 455
E1D	SpeI	1508 bp	715 + 793
E1E	EcoRV	1355 bp	783 + 572
E1F	ScaI	1664 bp	1313 + 351
E1G	DraI	1706 bp	848 + 250 + 608
E1H	ScaI	1455 bp	1112 + 343
E2	ScaI	1390 bp	401 + 989
ELB1	EcoRV	1014 bp	471 + 543
ELS1	SpeI	1535 bp	524 + 1011
ELB2	ScaI	1697 bp	157 + 1540
ELS2	SpeI	980 bp	286 + 694
ELB3	ScaI	1071 bp	728 + 343
ELS3	HindIII	1571 bp	638 + 933
ELS4	EcoRV	860 bp	535 + 345

* ‘ \approx ’ indicates the approximate size suspected by comparison with DNA marker. Because the E1N1 and E1N2 contain gap-sequence, the digested DNA fragments were different to expected fragments lengths.

Table III.3. List of primer pair sequences for 5' - and 3' - RACE-PCR		
Primer	Sequence (5' → 3')	T _m (°C)
RT primer	*p-ACAAGTAGCCCTCAG	50°C
5'-Race	1st round F: GTTGGGTTTTGTGGTTCGTT	54°C
	R: GTTGGAGAAAGAGGGGGACA	
	2nd round F: TGAGTGGACCTACGGCTT	56°C
	R: CCGATGGGCATAATACACA	
3'-Race	1st round F: GATATCAGCTGGATGCAGTTATTC	55°C
	R: CTGATCTAGAGGTACCGGATCC [†]	
	2nd round F: TCTTTCTTGAGGTGGGGGTG	60°C
	R: CTGATCTAGAGGTACCGGATCC [†]	

*RT-primer for 5'-RACE was phosphorylated at the 5'-end

[†]The reverse primer for 3'-RACE-PCR was kit supplied; both the 1st and 2nd round PCR was performed with the same reverse primer.

Table III.4. List of primer pairs for bisulfite sequencing					
Primer		Sequence (5' → 3')	Tm (°C)	Region [†]	Size (bp)
BS1	Outer	F: TGGTTAAATGAGGTATTTGGA R: CCATAAAACATAACTAAAACTAAA	54°C	-425 ~ +100	525 bp
	Inner (BS1-1)	F: TTTGTTATATTGTTTGTGGAAAA R: CCATAAAACATAACTAAAACTAAA	50°C	-329 ~ +100	429 bp
BS2	Outer	F: TTGGGATATTTTAAGGTAATTTT R: ATTTTATAAACATTCCAAACAATACA	56°C	+206 ~ +792	587 bp
	Inner 1 (BS2-1)	F: GGGATATTTTAAGGTAATTTT R: AAAAACAAATATCCATTACCCT	56°C	+208 ~ +557	350 bp
	Inner 2 (BS2-2)	F: TTTATTAGGGTAATGGATATTTG R: ATTTTATAAACATTCCAAACAATACA	56°C	+530 ~ +792	263 bp
BS3*		F: TTGTATTGTTTGGAATGTTTATA R: CCTTCTCCCTCTCATTTTCT	54°C	+766 ~ +1146	381 bp
BS4	Outer	F: AGAAAATGAGAGGGAGAAGGTT R: CCTTAAATACCACCCACTAAAAA	58°C	+1127 ~ 1795	634 bp
	Inner 1 (BS4-1)	F: AGAAAATGAGAGGGAGAAGGTT R: TATCCACATAACAAAATCAACCA	60°C	+1127 ~ +1452	326 bp
	Inner 2 (BS4-2)	F: AGGGATAATATGGTTGATTTTGT R: CCTTAAATACCACCCACTAAAAA	52°C	+1420 ~ +1759	341 bp

*BS3 was amplified without performing nested PCR.

[†]The region analyzed in this study spanned 2 kb upstream and downstream of one of the transcription start sites (289233rd nucleotide of the pig X-chromosome scaffold sequence, NW_003612825.1)

4. Results

Identification of a porcine *XIST* gene model and its RNA sequence

To identify the porcine *XIST* gene and its RNA sequence, pig genomic DNA sequence (X-chromosome scaffold, GenBank accession No. NW_003612825.1) was aligned and compared with mouse (NC_000086.7), human (NC_000023.10), and bovine *XIST* gene models (AC_000187.1) (Figure III.1). The DNA sequence ranging from nucleotides 289,233 to 257,103 in the pig scaffold was well-aligned with the three gene models. Specifically, they matched the TSS and the last nucleotide of the human and bovine *XIST* genes, respectively (Figure III.1 and Figure III.2A). Primer sets (Table III.1) were designed to confirm whether the aligned pig sequence (candidate *XIST*) is transcribed in PEF cells (Figure III.2A). RT-PCR using these primer sets confirmed that the candidate *XIST* gene was transcribed only in female PEFs, and the target regions were not detected in male PEFs (Figure III.3). Amplicons were also confirmed by restriction enzyme mapping, which produced patterns consistent with the predictions (Table III.2). All of the amplified regions were sequenced by Sanger sequencing, and most of the sequences were consistent with the porcine DNA sequence. However, one amplicon (E2EL) that partially covered the first and last exons of the candidate porcine *XIST* gene was shorter compared with the expected gDNA sequence

(Figure III.3), indicating the presence of introns and exons in the target DNA region. Two other amplicons (E1N1 and E1N2) exhibited longer bands than expected because of a gap (Ns) in the pig scaffold, which was filled with a 626-bp nucleotide sequence from the longer-than-expected amplicons (Figure III.4). Finally, the sequenced amplicons were assembled into a porcine *XIST* gene model consisting of two long exons (exon 1 and exon 7) and 5 smaller exons (exon 2 - exon 6) between exons 1 and 7 (Figure III.2B). The exons were located between introns with RNA-splicing recognition sequences (GT and AG, Table III.5) at their 5' and 3' termini, respectively. The results suggested that the porcine *XIST* gene shares an evolutionarily conserved region among mammals, indicating the reliability of model identified in this study.

Identification of the 5' and 3' end regions of porcine *XIST*

Transcript start sites (TSSs) of porcine *XIST* were identified by 5'-RACE PCR using the primer sets (Table III.3). The 289,233rd nucleotide of the pig X chromosome scaffold (NW_003612825.1) was expected to be a porcine *XIST* TSS based on sequence alignment with the human and bovine *XIST* 5' region (Figure III.1), which has been suggested to be conserved among mammals (Elisaphenko *et al.* 2008). Add to the expected site, 7 TSSs were searched additionally (Figure III.5). Two TSSs (-5 and -3) were found upstream of the expected TSS, while 5 TSSs (+16, +48, +72, +143, and +211) were located downstream of the expected

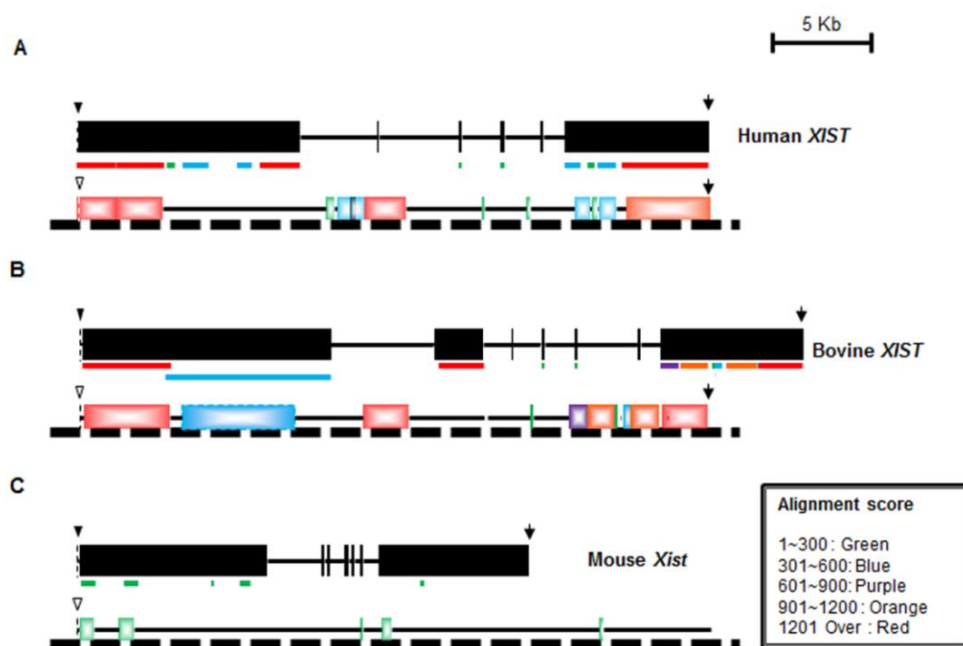
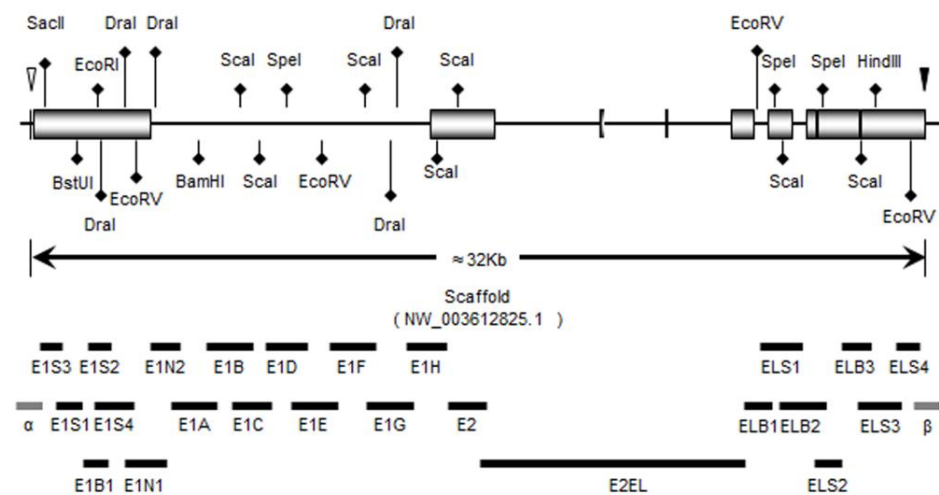


Figure III.1. *XIST/Xist* homologue analysis of the pig genome by BLAST search. (A) Human *XIST*, (B) bovine *XIST*, and (C) Mouse *Xist* (GenBank accession Nos. NC_000023.10, AC_000187.1, and NC_000086.7 respectively) models are represented as black boxes (exons) and black lines (introns). Black dashed lines under each *XIST/Xist* gene indicate the pig X-chromosome scaffold (NW_003612825.1). Filled arrowheads indicate the transcription start sites (TSS) of each *XIST/Xist* gene. Empty arrowheads present on the defined regions of the counterparts of *XIST/Xist* gene indicate candidate porcine *XIST* TSSs based on BLAST alignment (289233rd nucleotide of the NW_003612825.1 scaffold). Arrows indicate the terminus of the *XIST/Xist* gene and the candidate last-sequence of the porcine *XIST* gene identified by aligning human/bovine *XIST* RNAs to the pig genome sequence. Colored lines are homologue regions in each *XIST/Xist* gene, and the colored boxes are the counterparts of the lines. Each color represents an alignment score. The diagram is scaled.

A



B

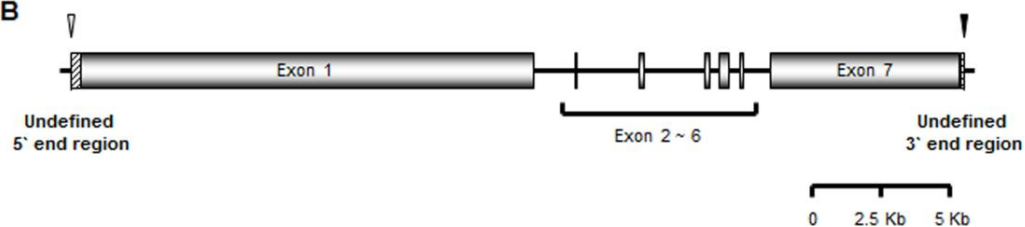
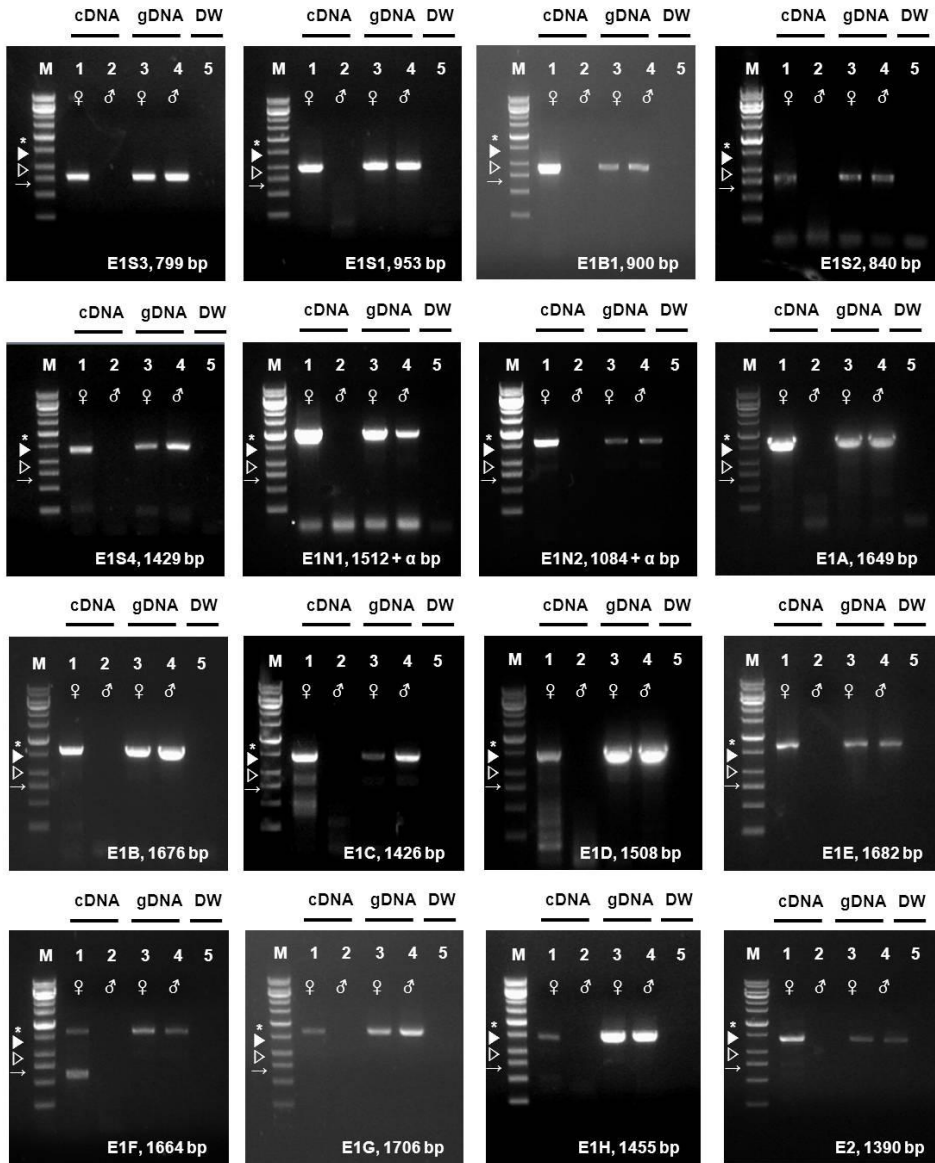


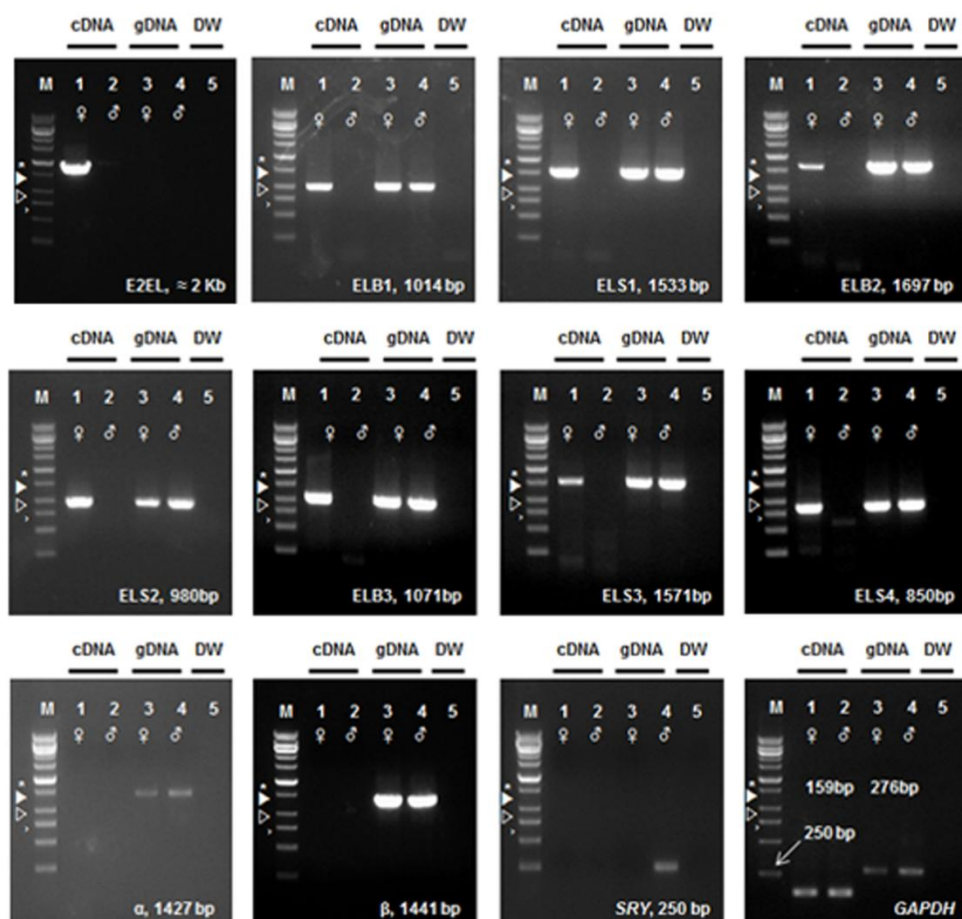
Figure III.2. Diagram of the candidate *XIST* expressing region in the pig. (A)

Designed *XIST* model and primer sets. Candidate *XIST* ncRNA expression regions on the pig X-chromosome scaffold (NW_003612825.1) determined by BLAST searches are represented by black boxes. Empty and filled arrowheads indicate the 289233rd and 257103rd nucleotides of the scaffold, which were aligned with the transcription start site (TSS) and the last sequence of human and bovine *XIST*, respectively. The designed primer sets are represented as black and gray lines: female-expressed primer sets are represented as black lines and regions that were not expressed in either sex (α and β) are represented as gray lines. Restriction enzymes used to digest each amplicon are shown. The diagram is scaled. (B) Diagram of the identified porcine *XIST* gene model. The porcine *XIST* expressing region was identified by sequencing. Filled rectangles indicate exons, and boxes with diagonal lines indicate undefined regions. Filled and empty arrowheads indicate candidate 5' and 3' end regions identified from alignment with human and bovine *XIST*, respectively. The diagram is scaled.

A



(Continued)



B

		α	E1S3	E1S1	E1B1	E1S2	E1S4	E1N1	E1N2	E1A	E1B	E1C	E1D	E1E
cDNA	Female													
	Male													
gDNA	Female													
	Male													
DW	No template													

		E1F	E1G	E1H	E2	E2EL	ELB1	ELS1	ELB2	ELS2	ELB3	ELS3	ELS4	β
cDNA	Female													
	Male													
gDNA	Female													
	Male													
DW	No template													

Figure III.3. RT-PCR analysis of each candidate region of *XIST* PCR amplicon expression. (A) Amplified PCR products of designed primer pairs. Female PEF cDNA (1), male PEF cDNA (2), female PEF gDNA (3) and male PEF gDNA (4) were used for amplification using each designed primer pairs and distilled water (5) was used as a negative control. The asterisk, filled arrowhead, empty arrowhead, and arrow represent the 2000, 1500, 1000, and 750 bp DNA bands of the ladder, respectively. (B) Summary heat map for PCR detection of porcine *XIST* RNA and genomic DNA sequences. Filled gray boxes indicate the presence of PCR target regions. White blank boxes indicate the absence of PCR target regions. The E2EL region was detected only in cDNA templates from female cells.

Figure III.4. Identified gap sequence in the pig X-chromosome scaffold, NW_003612825.1. A gap sequence region was amplified with primer pairs designed to span the gap region. An unknown region nearly 500 bp longer than the expected length (100 bp) was observed. The gray-boxed sequence (TTTAAA) indicates a DraI restriction site.

Table III.5. Exons and introns information of porcine *XIST*

Number	Sequence (5` to 3`)	Length (bp)
Exon 1	*TATTTCTTAGACTACT	17,184 bp
Intron1	** GTAAGTACTTTTAAAG	1,460 bp
Exon 2	GGATGAATCTCCAAAG	89 bp
Intron2	GTGAATCTTTCTCAAG	2,271 bp
Exon 3	GATATTCCAGAAAAAG	136 bp
Intron3	GTAATTTATTCTCCAG	2,136 bp
Exon 4	ATCTTCCTCCATCTGAG	209 bp
Intron4	GTGGGTAATCTTTTAG	356 bp
Exon 5	GAAAACAGTACTCTAG	329 bp
Intron5	GTCAGTGGTTCGGTAG	390 bp
Exon 6	CTCCTGATAGGATGAA	129 bp
Intron6	GTAAGTTGTCTTCCAG	989 bp
Exon 7	TGATTGTCAAAACCTTA	7,139 bp

*289,233rd nt of NW_003612825.1 scaffold sequence was set to first sequence which was one of the identified TSS and matched to blast search in this study.

**Bold sequence, GT/AG, indicate splicing recognition sequence

TSS. Transcription termination sites were also identified by 3'-RACE PCR. The 257,103rd nucleotide of the pig X chromosome scaffold (NW_003612825.1) was expected to be a porcine *XIST* termination site based on sequence alignment with the human and bovine *XIST* ncRNA sequence. One transcription termination site (257,094th nucleotide) was searched and it is located 10 bp downstream of the expected site. The 3' region was confirmed by RACE analysis, which indicated the presence of an AATAAA poly A-tail signal sequence located 21 bp upstream of the termination site (Figure III.5).

Repeat sequence analysis in porcine *XIST* RNA

The repeated sequences in *XIST/Xist* RNA have been suggested to have a function in X chromosome silencing. Therefore, it was examined whether these repeated sequences and conserved functional domains are present in the identified porcine *XIST* RNA sequence. Four repeat sequence regions were identified (Figure III.6A). Interestingly, the *Xist* A-repeats comprises copies of a 24-mer consensus sequence known to conserved between humans and mice (Brockdorff *et al.* 1992) were detected in the first repeat region of porcine *XIST*. The consensus sequence of the first repeat region was predicted to have a secondary structure containing two stem-loops (Figure III.6B), which is the same structure as that of mouse *Xist* (Wutz *et al.* 2002). The consensus 24-mer monomer was repeated 8 times in the +327 to +695 region (Figure III.6B).

A

```

-96 GGCTTTCAATCTTCTTGGCCACTCCTCTTCTACTCCCTCCGCCCTCAGTCCCCTCCCTCTCA -35
      *
-34 CTTCTTAAAGCGCTGCACTTTGCTGCGACCGCCATATTTCTTCTTTTTCCCGGGTGGAAGCT +28
+29 TGCTGGTATTGGATCTCTTTGCCCCGTGTGGTTCTTCTGGAACATTTTCCAGCCCCCAGCCA +90
+91 TGCCTTATGGCATATTTCTTTAAAAAAAATCCACCAAAAATTCATAAAATGTTTTAAAATT +152
+151 TCTAAACTTTCTCCTAATATTTTCTTGACACCTTATCTCTAGTTTACAGTTATTTGGGATAT +214
+215 TTTAAGGCAACTTTCTATTTTAAAATAATTTTCTTTGGAATGTTTTTTGG +265

```

B

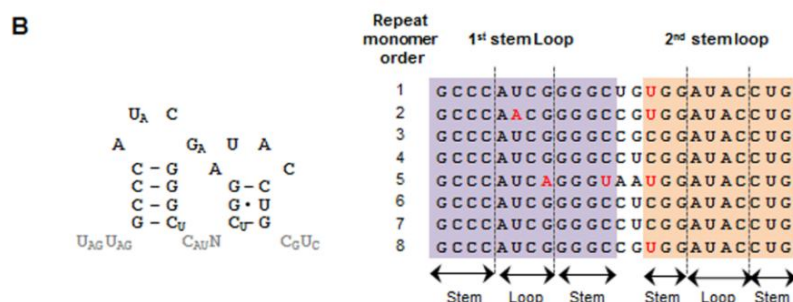
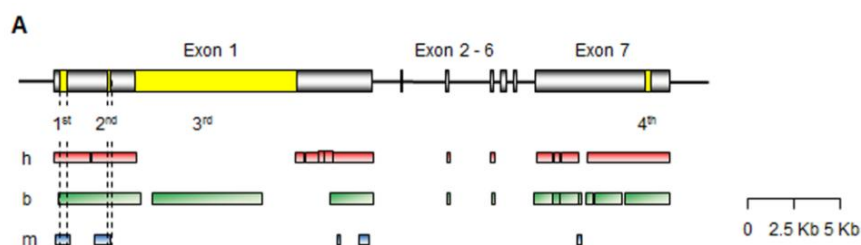
```

+32,761  ACAAATTGAAGAGTTTGTTGTTTAAATGTCAATAAAATACTGTTTT TTGAAAACCTT A +32,817
                                     ↑      ↑
                                257,103th nt 257,094th nt
( ScaffoldID : NW_003612825.1 )

```

Figure III.5. Identification of transcription start sites and the termination site of porcine *XIST*. (A) Transcription start sites (TSSs) of the porcine *XIST* gene were identified by 5' RACE-PCR. Eight putative TSSs were confirmed (bold characters), and the asterisk-marked sequence was determined to be the candidate TSS based on a BLAST homology search (289,233rd nucleotide of the pig X-chromosome scaffold sequence, NW_003612825.1). The TSS (289,223rd nucleotide of NW_003612825.1, asterisk-marked “T”) was set as +1. (B) The last sequence of porcine *XIST* was defined by 3' RACE-PCR. The italic sequence (“T”, 257,103rd nucleotide of NW_003612825.1) is expected to be the last sequence based on BLAST alignment. The sequence marked with an asterisk (“A”, 257094th nucleotide of NW_003612825.1 and 32,817th nucleotide from TSS) indicates the last sequence identified by 3' RACE-PCR. The gray-boxed region (AATAAA) represents one of the poly A-tail signal sequences.

The second repeat region contained a cytosine-rich 6-mer repeat that was located in the +2,723 to +2,870 region of the porcine *XIST* RNA as the cysteine-rich repeat-B region of mouse *Xist* (Brockdorff *et al.* 1992). In all, a total of 23 repeats were observed in this region (Figure III.6C). A consensus sequence of the mouse C-repeat sequence was also searched in porcine *XIST*, but did not locate such a sequence in porcine *XIST*. Two novel repeat regions (the third and fourth regions in Figure III.6A) were also identified. The third region comprised 91.7 copies of a 96-bp consensus sequence ranging from +4,302 to +13,102 in the porcine *XIST* gene (Figure III.6D). The fourth region consisted of two copies of a 149-bp monomer ranging from +24,059 to +24,356 (Figure III.6E). Interestingly, consensus sequences in the third region of porcine *XIST* were not present in either human or mouse repeat regions, while the fourth region was only partially aligned to a non-repeating region of human *XIST*. Taken together, the identified porcine *XIST* RNA contained not only conserved sequences, but also porcine-specific repeated sequences, suggesting conserved and hypothetically distinct functions of porcine *XIST*.



C

+2723	CCCCAG	CCCCUG	CCCCUG	CCCCUG	CCCCAG	CCCCUG	CCCCAG	CCCCUG	CCCCUG	+2782
+2783	CCCCAG	CCCCAG	CCCCUG	CCCCUG	CCCCUG	CCCCAG	CCCCAG	CCCCAG	CCCCAG	+2841
+2842	CCCCAG	CCCCAG	CCCCAG	CCCCAG	CCCCAG	CCCCAG	CCCCAG	CCCCAG	CCCCAG	+2870

D

Consensus sequence TTAATCTAGGATCCCATCCCCACCTAATCCATTTCATTACTGCAGAGGGTACATGGGAC-TGTTAATGGGCTTAAATGCAGCTACAGTCCCTTATG

1st rep. TTAACTAGAGTCTCACTCA-ATAACCGCTTTGCATAAGAGTAGAGGATCTTGTGAC-AGTTAAGTGGTTTGAITGCAGCTACAGACCTGCTG

2nd rep. TTAGACTAGCATCCATCTCTCTCTAGAT-GATTACATGACCAATCCGTACTAGGGAC-TGTTAATGGGTTAAATGCAGCTACAGTCCCTTATG

3rd rep. TTAATCTAGGATCCCATCCCCAATCACTCTCTTTTCACTACTGCAGGGCTTACATAGGACTTGTAAAGACCTTTATTGCA-CTACTGTCGCTTCTG

4th rep. TTAGTCCAGACCCCATATCTCTCTAAATCC---TGTATTACCG-AAAGAAATATAAGGGAC-TGTTAAGTGTCTAAA-GCAGTTACAGTCCCTTATG

5th rep. TTTATCTAGGA-CCCATCTCTCACTAACCCCTTTTGCATGACTGCAGAGGGTCTCTGGGAC-TGTTAATGGGCTTTATTGCAGCTACAGTCCCTTCTA

92nd rep. CT-AT-TAGTTTCCAATCCCTTTCT-ACTCCCTTTGCATTACCAAGAGGGTACTTGGGAC-TGTT

E

+24, 059	CAGAGAGGTGGATAGGAGGAAGTGAAGAAAAGGTGAGACCTAGTTCGATGGGCAACCTGCCTTTGTCTG	+24, 133
+24, 208	CAGAGAGGTGGATAGGAGGAAGTGAAGAAAAGGTGAGACCTAGTTCGATGGGCAACCTGCCTTTGTCTG	+24, 282
+24, 134	GATTGGTCTTAAAGTATTACTAOCCTCAGATCTAACTTGGGGATTGCACTCTTAGTTTAAATGCAACTCCAG	+24, 207
+24, 283	GATTGGTCTTAAAGTATTACTAOCCTCAGATCTAACTTGGGGATTGCACTCTTAGTTTAAATGCAACTCCAG	+24, 356

Figure III.6. Analysis of repeat sequences in porcine *XIST* RNA. (A)

Distribution of repeated sequences in the porcine *XIST* gene model. Gray rectangles indicate exons, and the yellow boxes in the exons indicate repeat regions. The regions having homology with human, bovine, and mouse *XIST/Xist* are represented to red (h), green (b), and blue (m) boxes, respectively. The first and second repeat regions showed similarity with all accessed *XIST/Xist* sequence from other species (dashed lines). The diagram is scaled. (B) The first repeat region in porcine *XIST*. A monomer of the first repeat region in pigs (left panel) share a consensus sequence and a predicted secondary structure of the mouse (Wutz *et al.* 2002). Gray text indicates non-conserved sequences. Eight repeated monomers identified in the porcine *XIST* first repeat region (+327 to +695) were aligned with a consensus sequence (Right panel). Red character indicates different sequence compared to the consensus sequence in conserved stem-loop region. (C) The second repeat region (+2,723 to +2,870) is shown, with non-repeat sequences in gray. (D) The third repeat region (+4,302 to +13,102) fraction. The first five repeats and last repeat were aligned with the consensus sequence. Red characters mean the sequences which were not same to consensus sequence. (E) Alignment of the fourth repeat region (+24,059 to +24,356). Two copies of the monomer were perfectly matched.

Expression of *XIST* in other porcine tissues

I assembled a full-length porcine *XIST* RNA sequence by PCR and sequencing and identified its 5' and 3' regions. Although assembly and identification of both ends of the RNA strand was performed using PEFs, it was not immediately clear whether *XIST* was also expressed in other porcine tissues in the same manner as gene model suggested by PCR and sequencing. Thus, to confirm identified gene model and verify the expression of *XIST* in other tissues, Illumina sequence reads from three different porcine tissues—liver, abdominal fat, and the longissimus dorsi muscle—of two female pigs were aligned with identified porcine *XIST* gene model (Figure III.7A). The Illumina sequence reads were well aligned onto the most of the region of the model, including the 5' and 3' regions (Figure III.7), while none of the Illumina reads were aligned to the upstream/downstream regions of the assembled *XIST* RNA (Figure III.7B). Thus, these results suggested that *XIST* is expressed in other porcine tissues as described above.

Analysis of methyl CpG sites in the upstream region of porcine *XIST*

After identifying the TSSs of the porcine *XIST* gene, the methylation status of the CpG sites upstream of the *XIST* TSS was analyzed. The search for CG dinucleotides was conducted using the ± 2 kb region of the first porcine *XIST* sequence (289,233rd nucleotide of NW_003612825.1, Figure III.8A), which

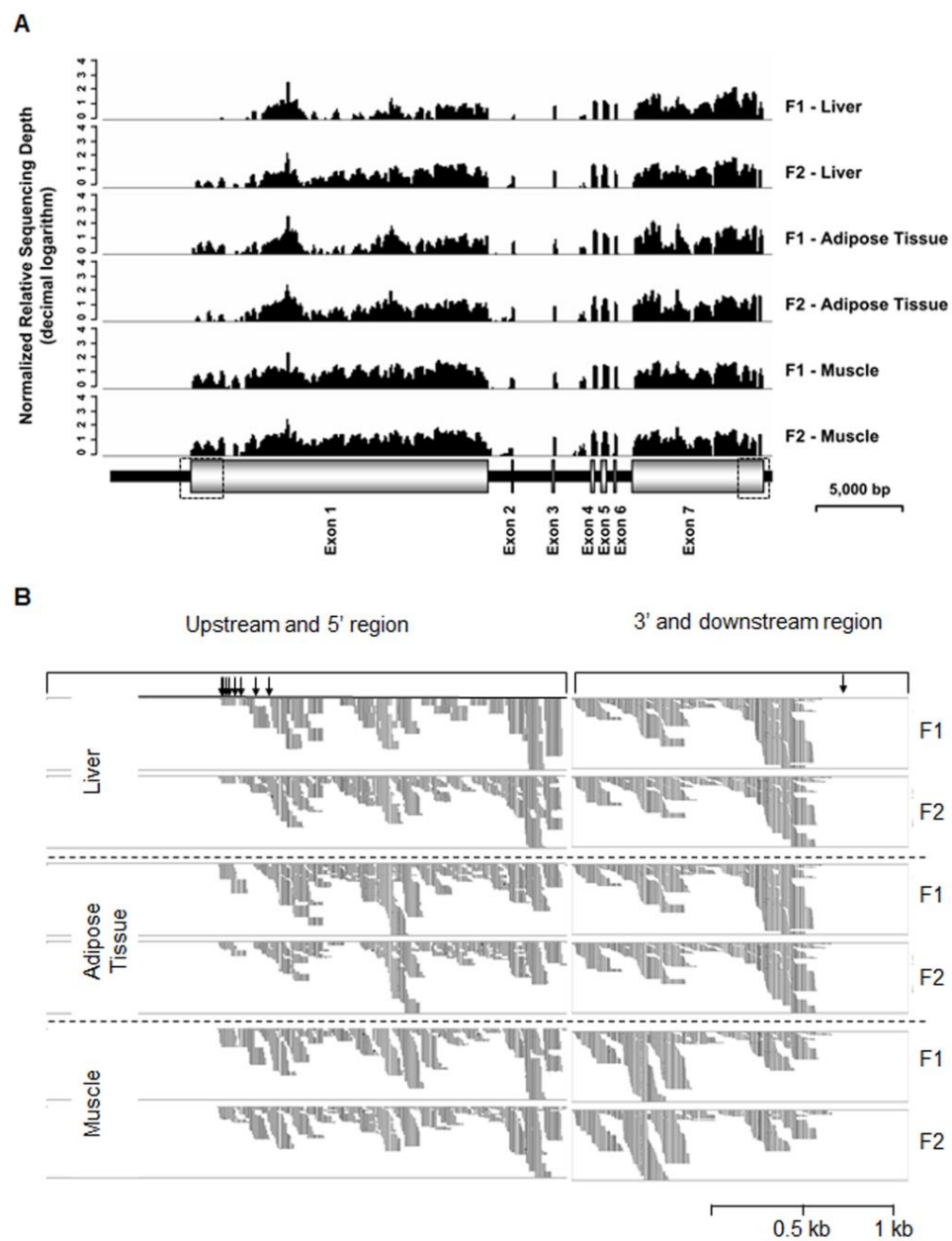
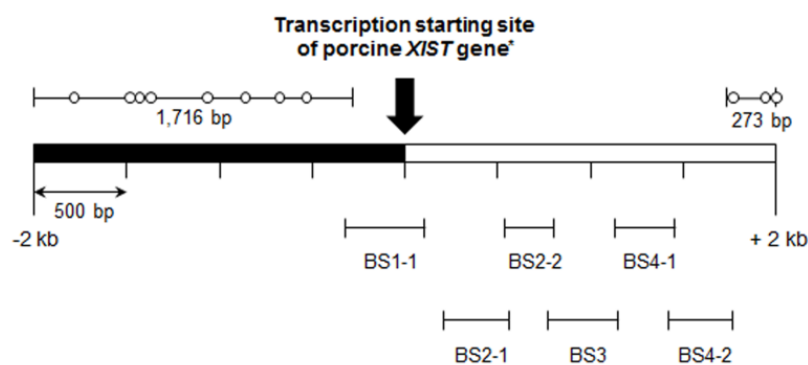


Figure III.7. Alignment of Illumina reads to the porcine *XIST* gene model. (A)

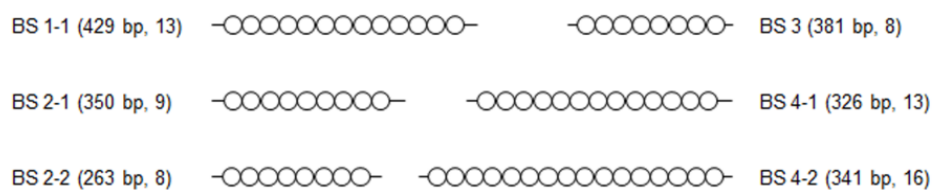
Alignment of Illumina reads to the porcine *XIST* gene locus. Six Illumina reads from three tissues (liver, abdominal fat, and longissimus dorsi muscle) of two female pigs (F1 and F2) were aligned to the entire porcine *XIST* gene including introns and exons. Small peaks between exon 3 and 4 were identified to be non-specific alignment of reads. The diagram is scaled. The two dotted boxes indicate upstream/5' and 3'/downstream regions of the *XIST* gene, respectively. (B) Alignment of Illumina reads to upstream/5' and 3'/downstream regions. Arrows in the upstream and 5' region indicate the transcription start sites (TSS) of the porcine *XIST* gene identified by 5' RACE. The arrow at the 3' and downstream region indicates the end of the porcine *XIST* RNA identified by 3' RACE.

identified 78 CG dinucleotides, 67 of which were densely located between nucleotides -284 and +1,727 (Figure III.8A). The methylation statuses of the 67 CpG sites in both male and female embryonic fibroblasts were analyzed using six different primer sets (Table III.4, and Figure III.8B). With the exception of five CpG sites, the analyzed regions were over 90% methylated in male PEFs, and the total methylation rate for male PEFs was 96.26% (Figure III.8C). However, in female PEFs, although five CpG sites on the BS1-1 region were highly methylated (>70%), the remaining CpG sites were methylated at a rate of 40~60%. Thus, the total methylation rate of the 67 CpG sites in female PEFs was 49.85% (Figure III.8C). The results suggested the analyzed CpG sites were differentially methylated between the male and female.

A



B



C

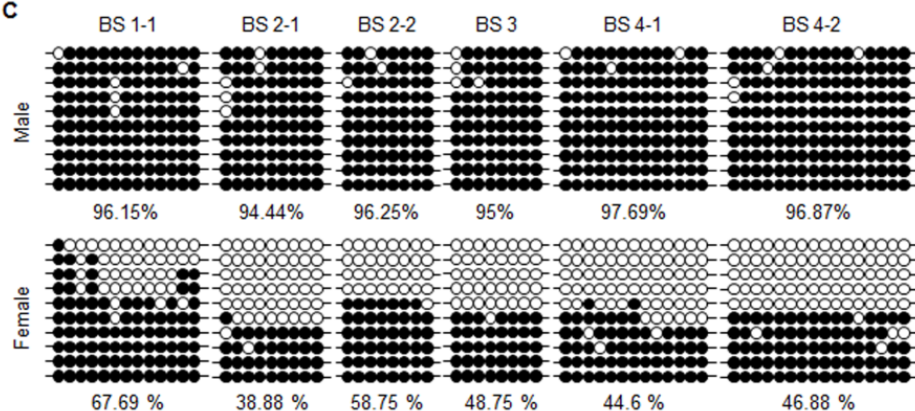


Figure III.8. Diagram of the CpG sites located in the ± 2 kb region of the transcription start site of porcine *XIST* and the methylation status of *XIST* CpG sites in male and female porcine embryonic fibroblasts. (A) One of the transcription start sites (TSS) identified by 5'-RACE-PCR was set as a +1 (asterisk, 289233rd nucleotide of NW_003612825.1). Black and empty bars indicate regions upstream and downstream of the TSS, respectively. Methylation of CpG sites in six regions (BS1-1, BS2-1, BS2-2, BS3, BS4-1, and BS4-2) was analyzed. The diagram is scaled. (B) Profiles of the six target regions. Each circle represents a single CG dinucleotide identified in the amplified region. The length of each amplicon was 250 bp – 450 bp and contained between 8 - 16 CG dinucleotides. The diagram is not scaled. (C) The methylation status of 67 *XIST* CpG sites were analyzed in male and female embryonic fibroblasts. Each circle indicates a CpG dinucleotide. Filled and empty circles represent methylated and unmethylated CpG dinucleotides, respectively. The horizontal line represents one individual clone.

5. Discussion

To the best of my knowledge, this study is the first to report a gene model of porcine *XIST* that includes a ncRNA sequence and the CpG methylation status of its regulatory regions. The purpose of XCI is to balance the expression level of X-chromosome-linked genes between male and female eutherian mammals, occurs early during embryo development (Lyon 1961; Okamoto and Heard 2009; Wutz 2011). The initiation of XCI is induced by cis-binding of *XIST/Xist* onto the X chromosome (Penny *et al.* 1996; Wutz and Jaenisch 2000). Interestingly, among the different strategies for XCI in eutherian mammals (Okamoto and Heard 2009; Okamoto *et al.* 2011), all appear to use *XIST* as a trigger for XCI. The importance of *XIST/Xist* has been a focus of not only XCI research, but also, more recently, stem cell research (Navarro *et al.* 2008; Nichols and Smith 2009) and embryo cloning research (Inoue *et al.* 2010; Matoba *et al.* 2011). Previous studies have indicated that *XIST* is associated with stem cell differentiation (Navarro *et al.* 2008) and abnormalities of cloned embryos in the mouse (Inoue *et al.* 2010). Despite the importance of *XIST* and XCI in a number of research fields, only a partial sequence of porcine *XIST* has been previously reported (Cepica *et al.* 2006; Groenen *et al.* 2012; Bischoff *et al.* 2013). Add to this, the porcine *XIST* expression and methylation status on its candidate CpG sites were analyzed in one previous report

comparing porcine embryonic stem cells from different origins and induced pluripotent stem cells. However, the information about porcine *XIST* used for the study was restricted and kinds of putative ones, so identification of porcine *XIST* whole sequence and analysis of its regulating sites should be accessed to perform more accurate and detailed *XIST* analysis in pig pluripotent cells and embryos. This led us to generate a porcine *XIST* gene model, including its ncRNA sequence, and to examine the CpG methylation status of its regulatory regions.

Sequence alignment was used to identify candidate porcine *XIST* exons, which showed not only an evolutionarily conserved *XIST* gene model among mammals, but also species-specific sequence divergence. Several evolutionarily conserved regions, such as repeated sequences among mammalian *XIST* orthologs, have been previously reported (Brockdorff *et al.* 1992; Elisaphenko *et al.* 2008). Indeed, the A-F regions are regarded as functional domains of *Xist* (Brockdorff *et al.* 1992; Wutz *et al.* 2002; Sarma *et al.* 2010), some of which were present in identified model of porcine *XIST*. On the other hand, divergence of *XIST/Xist* sequences among species due to rapid ncRNA evolution has been noted (Hendrich *et al.* 1993; Pang *et al.* 2006). Consistent with this observation, porcine *XIST* also exhibited sequence divergence in some regions compared with other mammalian *XIST/Xist* genes, and showed varying sequence similarity depending on the compared target species. However, certain regions, such as the 5' region of exon 1, were highly conserved among species, as shown in Figure III.1.

A previous study suggested that *XIST/Xist* originated from a protein-coding gene (*Ln timer*) and transposable elements (Elisaphenko *et al.* 2008), and specifically that the 5' region of exon 1 and exons 5 through 7 of the consensus *XIST* sequence arose from a protein-coding gene. In this study, it was shown that the same regions of human and bovine *XIST* were well-aligned to that of porcine *XIST* (Figure III.1), suggesting that the highly conserved domains originating from the proposed protein-coding gene survived in pigs during evolution, and may have a role in porcine *XIST* function. In contrast, some regions originating from transposable elements are suggested to have integrated into *XIST* after mammalian taxa divergence, resulting in species-specific sequence variation among eutherian *XIST/Xist* (Elisaphenko *et al.* 2008). Such sequence variation may account for the presence of porcine-specific regions that did not align to the *XIST/Xist* sequences of other species, and which may also contribute to species-specific modes of *XIST* function during XCI. Together, these evolutionarily properties of *XIST/Xist* support the reliability of the porcine *XIST* expression region.

Most of the RT-PCR target regions, except the α and β regions, were amplified exclusively in female porcine cells (Figure III.2 and III.3), which was highly consistent with previous observations in other mammals (Chow and Heard 2009; Okamoto and Heard 2009; Augui *et al.* 2011; Wutz 2011). In particular, primer pairs that amplified regions that were not selected as part of the first candidate model (E1N1 to E1H in Figure III.2A) exhibited female-specific expression (Figure III.3). These regions may have evolved in a porcine-specific manner as

suggested previously (Elisaphenko *et al.* 2008), and may contribute to a novel role of *XIST*-associated XCI distinct from that in other species.

In the mammalian *XIST/Xist* gene model, the first and last exons are much longer than the central exons (Brockdorff *et al.* 1992; Brown *et al.* 1992). As expected, the E2EL amplicon contained 5 small exons between the first and last exons, consistent with other species (Figure III.2B). Furthermore, the identified exon 4 was conserved in pigs (Figure III.1 and (Yen *et al.* 2007)). These results suggested that the porcine *XIST* ncRNA shares a common model with other mammalian *XIST/Xist* genes.

The identified porcine *XIST* gene model covered a previously annotated partial porcine *XIST* sequence in the 3' region of *XIST* (GenBank accession Nos. AJ429140 and EF619477.1). In addition, the recently annotated partial exon 1 (nannotator00001430) and exon 4 (nannotator00001519) sequences were also located in the expected porcine *XIST* exons (<http://rth.dk/resources/rnannotator/susscr102/>). The presence of a previously reported partial porcine *XIST* expression region on the defined *XIST* gene supports the reliability of porcine *XIST* confirmed in here.

The repeat sequences in *XIST/Xist* have been suggested to have functional roles in XCI (Brockdorff *et al.* 1992; Brown *et al.* 1992). Interestingly, some repeat regions were detected in the porcine *XIST* ncRNA sequence. The A repeat, which consists of 7.5 repeats of a 24-bp sequence, forms two stem-loop in mouse, and has

been known to regulate processes essential for X chromosome silencing (Wutz *et al.* 2002). Interestingly, Wutz and colleagues showed that a mutated A region allele results in a failure to induce X chromosome silencing in a stem cell model without affecting *Xist* stability and localization (Wutz *et al.* 2002). The identified porcine *XIST* 5' region included a total of 8 copies of the 24-bp sequence (Figure III.6A and III.6B), and this region was well-aligned to the 5' region of mouse *Xist* (Figure III.6A). This result suggested that, similar to the A-repeat in mice, the conserved first repeat region in porcine *XIST* may be involved in silencing X chromosomes, although this possibility needs to be further analyzed. Another repeat region, the C-repeat, which consists of 14 copies of an approximately 100-bp sequence in mice, has been suggested to be involved in *Xist* localization during XCI (Sarma *et al.* 2010). However, targeting of the C region with Locked nucleic acids (LNAs) does not affect the localization of human *XIST* (Sarma *et al.* 2010), which has only one copy of the C-repeat (Brockdorff *et al.* 1992; Yen *et al.* 2007). The porcine *XIST* sequence has a sequence similar to that of the human C-repeat sequence (49 bp downstream of the second repeat region), while only one copy was found in the pig. Thus, these results may suggest that porcine *XIST* has a novel mechanism of regulating *XIST* localization on the X chromosome compared with the mouse. Two other repeat regions of porcine *XIST* (regions 3 and 4 in Figure III.6A) were identified as candidate functional domains. In particular, the third repeat region consisted of a highly repeated sequence (91.7 copies, Figure III.6D) and was found to be well-aligned only with the bovine *XIST* sequence (Figure III.6A). The D-

repeat has been suggested to be the largest repeat, and is commonly found in exon 1 of *XIST/Xist* in many species (Yen *et al.* 2007; Elisaphenko *et al.* 2008). The D-repeat is hyper-variable with respect to both its sequence and length among species (Yen *et al.* 2007). Thus, it was hypothesized that the third repeat in porcine *XIST* may be either a D region or a porcine-specific repeat. While four repeat regions were identified in porcine *XIST*, their biological roles need to be clarified in future studies.

The alignment of Illumina reads (NCBI SRA accession numbers: SRX054582 - SRX054587) with newly suggested *XIST* gene model clearly demonstrated that the identified regions were expressed in other tissues according to the same model and that transcription start and termination sites at the 5' and 3' regions, respectively, were consistent among different porcine cell types (Figure III.7). Taken together, the results suggest that new model of the porcine *XIST* gene and the estimated 5'/3' regions are reliable.

The promoter of *XIST/Xist* is important for regulating differential expression between female and male eutherians. Two representative *XIST* promoters, P1 and P2, have been reported. P1, a minimal promoter, is located close to the TSS (Pillet *et al.* 1995; Hendrich *et al.* 1997), while P2 is located nearly 1 kb from P1 (Johnston *et al.* 1998; Elisaphenko *et al.* 2008). In present study, there was a region of high CG dinucleotide density from the -284th nucleotide to +53rd nucleotide and from the +285th nucleotide to +1727th nucleotide. The distribution patterns of CpG

sites of porcine *XIST* were similar to P1 (Pillet *et al.* 1995; Hendrich *et al.* 1997) and P2 (Johnston *et al.* 1998; Elisaphenko *et al.* 2008) of other species. A previous study demonstrated that the *Xist* promoter is highly methylated on active X chromosomes in somatic cells, while inactive X chromosomes do not have a methylated promoter region (Norris *et al.* 1994). Likewise, male porcine somatic cells carrying an active X chromosome were also highly methylated at *XIST* regulatory regions, while female porcine somatic cells carrying both active and inactive X chromosomes displayed a half methylation pattern (Figure III.8C). Results in this study were consistent with those of previous reports, and support the hypothesis that the analyzed porcine CpG sites may be regulatory regions of *XIST* expression. Several transcription factors, such as CTCF, YY1, and TBP1, are known to regulate *XIST/Xist* expression at promoter regions (Pillet *et al.* 1995; Hendrich *et al.* 1997; Sheardown *et al.* 1997b; Pugacheva *et al.* 2005). Interestingly, a TBP1 consensus sequence was confirmed in the candidate porcine P1 region (TTAAAG, -30 to -25), suggesting that TBP1 may be a regulator of porcine *XIST* expression. The presence of this TBP1 consensus sequence in 30bp upstream of one of the TSS, the 289,233rd nucleotide of the pig X chromosome scaffold (NW_003612825.1), suggest that the TSS which was applied to the first sequence of porcine *XIST* in this study is most reliable TSS of the porcine *XIST* RNA. Although CTCF-binding sites in *XIST* gene and its upstream region were tried to find, because of its diverse binding sites, it was hard to determine presence of suitable sites for CTCF binding. The YY1 of which consensus sequences,

CCGCCATNTT, is also present in near the *XIST* TSS (CCGCCATATT, -6 to +4), but it is unclear this region is actually recruited as an YY1 binding and functional site. So it should be analyzed these transcription factors are also functional in porcine *XIST* expression as in other species *XIST/Xist* expression. So the exact porcine *XIST* promoter regions also need to be further defined in order to clarify the relationship between CpG methylation and transcription factor binding on *XIST* expression using like chromatin immunoprecipitation (ChIP) - assay.

In this study, the entire coding region of porcine *XIST* was identified using sequence alignment, RT-PCR, Sanger sequencing, restriction enzyme mapping, and Illumina alignment. I proposed a porcine *XIST* ncRNA 25,215 bp in length and comprising seven exons (Table III.5, GenBank accession No. KC753464). CpG methylation in the regulatory regions of porcine *XIST* was also examined. The porcine *XIST* gene model and CpG methylation profile reported here may be important for understanding the mechanism of XCI in pigs. Further studies will be needed to determine the implications of *XIST* model identified in this study on the stem cell biology and embryology of pigs.

CHAPTER IV

**Some XIC-linked genes achieve dosage
compensation in porcine blastocysts**

1. Abstract

X-chromosome inactivation (XCI) is an epigenetic mechanism that occurs in the eutherian embryo development with the purpose of equalizing the dosage of X-linked genes between males and females. This event is regulated by various factors, and the genes located in the X-chromosome inactivation center (XIC), which is known to be an evolutionary conserved region, are associated with XCI. However, a number of studies regarding this epigenetic event and genomic region are primarily performed in mouse models despite its species-specific features. Thus in this study, the porcine XIC was identified and expression of the genes linked in the region was analyzed in porcine preimplantation embryos. Comparative sequence analysis revealed that the porcine and human XIC have conserved synteny and the non-coding RNAs were less conserved compared to the protein coding genes in the XIC. Among XIC-linked genes, expression levels of *CHIC1* and *RLIM* were decreased in morula to blastocyst development and their dosage was compensated between male and female blastocysts. Also, CpG sites of *CHIC1* were half methylated in female blastocysts. Contrary to these genes, *XIST* and *LOC102165544*, an uncharacterized non-coding gene, showed a dramatic increased expression levels after the morula stage and showed female preferential expression in blastocysts. Also, their CpG sites were hypo-methylated in female blastocysts. These results demonstrate that the porcine XIC consist of an evolutionary

conserved structure with fewer sequence conserved non-coding RNAs. In addition, it is likely that a few XIC-linked genes would achieve dosage compensation but chromosome-wide inactivation of X-linked genes is not accomplished in porcine blastocysts.

Key words: Dosage compensation, X-chromosome inactivation center, X-chromosome inactivation, preimplantation embryo

2. Introduction

X-chromosome inactivation (XCI) is an essential event for balancing the dosages of X-linked genes between male and female eutherians (Lyon 1961; Lyon 1962). This process is induced by the X-chromosome inactivation specific transcript (*XIST*) (Penny *et al.* 1996). Prior to the identification of this non-coding RNA (ncRNA), one report suggested that a specific region in the X-chromosome is required for inactivation of one of the X-chromosomes in mice (Rastan 1983). This region has been called the X-chromosome inactivation center (XIC) and various protein-coding and non-coding genes including *XIST* are present in this region. Protein-coding genes in the XIC have been considered to be conserved among non-eutherian vertebrates as well as eutherians. However, orthologs of the ncRNA genes in the XIC, like *XIST*, were not discovered in non-eutherians vertebrates, and the genes are less conserved among the eutherians. This is due to a disruption of the ancestral protein-coding genes and their pseudogenization resulted in evolving non-translated RNA genes in marsupials during an evolutionary short period (Romito and Rougeulle 2011). The ncRNAs in the XIC, *Tsix*, *Jpx*, and *Ftx* as well as *Xist*, were reported to regulate the XCI in mice (Sado *et al.* 2005; Tian *et al.* 2010; Chureau *et al.* 2011). The XIC-linked protein-coding gene, *Rlim* (also known as *Rnf12*), has also been reported to induce *Xist* expression in mice (Jonkers *et al.* 2009). These reports suggest that the XIC-linked genes are important for regulating

XCI in eutherians. Although the porcine *XIST* gene was first identified in previous study (Hwang *et al.* 2013a), little is known about porcine XIC

A previous report suggested that there is diversity in XCI initiation strategies and mechanisms to achieve this epigenetic phenomenon among the eutherians (Okamoto *et al.* 2011). Chromatin status of mouse X-chromosome is undergone serially changed during early embryo development like imprinting XCI of the paternal X-chromosome (Xp) in cleaved embryo, reactivation of inactive Xp in inner cell mass (ICM), and random XCI in differentiating ICM during early embryonic stages. (Augui *et al.* 2011). However, XCI in human and rabbit embryos differ to those in mouse embryo. Embryos of both species aren't undergone imprinting XCI and express *XIST* bi-allelically. In addition, initiation timing for XCI is delayed compared to the random XCI in mice. (Okamoto *et al.* 2011). The report suggests the requirement of species-specific studies on XCI mechanisms in various eutherians.

Compared to the studies regarding human and mouse XCI, which were mainly performed by cytological analysis using RNA fluorescence *in situ* hybridization (FISH), a few studies were have been performed to determine the XCI status in ungulates by comparing the expression levels of X-linked genes between males and females (Bermejo-Alvarez *et al.* 2010; Bermejo-Alvarez *et al.* 2011; Park *et al.* 2012). Global gene expression analysis demonstrated that sexual dimorphic expression was observed in bovine blastocysts and more than 80% of the X-linked

genes that are expressed in the blastocyst stage were expressed preferentially in female blastocysts rather than male blastocysts (Bermejo-Alvarez *et al.* 2010). Also, the expression levels of the selected genes, which were reported to be expressed differentially by sex in blastocysts (Bermejo-Alvarez *et al.* 2010), were equalized between males and females in elongated bovine embryos (Bermejo-Alvarez *et al.* 2011). Although one previous study observed sexual dimorphism of selected X-linked genes in porcine blastocysts (Park *et al.* 2012), more studies are required to confirm the XCI status in porcine embryos.

In this study, the porcine XIC was identified by sequence comparison with human and mouse genes and the expression of the genes in the porcine XIC was analyzed in the embryonic stage and in blastocysts to confirm the dosage compensation status in pig embryos. Finally, the methylation patterns of the promoter region of each XIC-linked gene were examined in female blastocysts. From this result, it was concluded that the dosage of a few XIC-linked genes was compensated, but chromosome wide inactivation of X-linked gene is not complete in female porcine blastocysts.

3. Materials and Methods

Ethics statement

All experiments are conducted with the approval of the Institutional Animal Care and Use Committees of Seoul National University (SNU-140325-3).

Sequence homology analysis

The genomic and transcript sequences of human and mouse XIC-linked genes were compared to porcine X-chromosome sequences (*Sus scrofa* version 10.2) using BLAST. The genomic range between sequence homologue regions with the 5' and 3' regions of each human and mouse XIC-linked gene was also used. Each genomic or transcript sequence of the human and mouse XIC-linked genes was compared to its counterpart region in pig, and their sequence homology was calculated.

Sequence homology analysis

Query sequence (sequence of the human or mouse XIC-linked gene) was aligned to the pig genome using BLAST. The query coverage of individual blast hits (A)

was calculated first as follows:

$$\text{Query coverage (A)} = \frac{\text{Length of each blast hit in query}}{\text{Query Length}}$$

The sequence homology rate of each blast hit (B) from alignment of one query was calculated by magnifying the sequence identity between blast hits and their searched counterpart region in pig:

$$\begin{aligned} &\text{Homology rate of individual blast hit (B)} \\ &= \text{Query coverage of blast hit (A)} \times \text{Identity} \end{aligned}$$

After calculating the homology rate of each blast hit (B) from the alignment of one query, their sum was considered to be the homology rate of the query (C), which is a partial fragment of the genomic sequence or transcript sequence of human and mouse XIC-linked genes.

$$\begin{aligned} &\text{Homology rate of query (C)} \\ &= \text{Sum of Homology rate of individual blast hit (B)} \end{aligned}$$

The homology rate of XIC-linked genes (D) of which partial fragment sequences were used for query, was calculated by averaging the homology rate of each query.

$$\textit{Homology rate of gene (D)} = \textit{Average of Homology rate of query (C)}$$

Sample preparation

Embryos produced *in vitro* To prepare *in vitro* fertilized and parthenogenic embryos, oocyte maturation was conducted first, following the protocol in previous report (Hwang *et al.* 2013b). Briefly, ovaries from pre-pubertal gilt were donated by the Sooam Biotech Research Institute (Seoul, Korea). Cumulus-oocyte-complexes (COCs) with granulated cytoplasm and thick-layered cumulus cells were collected from the porcine follicular fluid (pFF). Collected COCs were washed using TL-Hepes-PVA medium (Funahashi *et al.* 1997) and cultured in tissue culture medium (TCM-199; Life Technology, Rockville, MD) supplemented with 10 ng/ml epidermal growth factor (EGF), 1 µg/ml insulin (Sigma-Aldrich, St. Louis, MO, USA) and 10% pFF for 44 hours at 39°C in 5% CO₂ conditions. The hormones, 4 IU/ml of hCG and eCG (Intervet, Boxmeer, Netherlands), were treated for only the first 22 hours. Cumulus cells were detached using 0.1% hyaluronidase (Sigma-Aldrich, St. Louis, MO, USA). Denuded oocytes were assessed for *in vitro* fertilization and parthenogenesis. To produce fertilized

embryos, ejaculated sperm purchased from the DARBI A.I. Center (Jochiwon, Korea) was washed with Dulbecco's phosphate buffered saline (DPBS; Welgene, Seoul, Korea) containing 0.1% bovine serum albumin (BSA; Sigma-Aldrich, St. Louis, MO, USA) by centrifuge. Cumulus-free oocytes and washed sperm with the final concentration of 1×10^5 cells/ml were placed together in 50 μ l of modified tris-buffered medium (mTBM) (Abeydeera and Day 1997) and co-incubated at 39°C in 5% CO₂. After 6 hours of incubation, the oocytes were washed twice using DPBS (Welgene, Seoul, Korea) with 0.1% BSA and transferred to porcine zygote medium 3 (PZM3) (Yoshioka *et al.* 2002). Parthenotes were generated by activating denuded oocytes through electric pulse (1.0 kV/cm for 60 μ sec) using the BTX Electro-cell Manipulator (BTX, CA) in activation medium (280 mM mannitol, 0.01 mM CaCl₂, and 0.05 mM MgCl₂). The oocytes were transferred and incubated with PZM3 containing 2 mM 6-dimethylainopurine (Sigma-Aldrich, St. Louis, MO, USA) for 4 hours, followed by PZM3.

Embryonic fibroblasts Male and female PEF cell lines from 27-day-post-coitus Yucatan mini-pigs were provided by the Sooam Biotech Research Institute (Seoul, Korea). PEFs were cultured in high glucose Dulbecco's modified eagle medium (DMEM; Welgene, Seoul, Korea) containing 10% fetal bovine serum (FBS; J. R. Scientific, Woodland, CA), 1% penicillin/streptomycin, 0.1 mM 2-mercaptoethanol, and 1.7 mM L-glutamine (Invitrogen, Carlsbad, CA, USA).

RNA extraction and reverse transcription

The zona pellucida of embryo was removed by Tyrode's acid (Sigma-Aldrich, St. Louis, MO, USA) and the RNA was extracted using the Dynabeads[®] mRNA DIRECT[™] Kit (Invitrogen, Carlsbad, CA, USA) in accordance with manufacturer's instrument. Pooled mature oocytes (n=40), parthenotes in each stage (2-cell, n=20; 4-cell, n=20; 8-cell, n=20; morula, n=10; blastocyst, n=5), and individual parthenogenic and fertilized blastocysts were assessed. Total RNA from PEFs (5×10^5 cells) was extracted using Trizol[®] Reagent (Invitrogen, Carlsbad, CA, USA) and then treated with DNase (Turbo[™] DNase I, Applied Biosystem, Foster City, CA, USA) following the manufacturer's instructions. One microgram of DNase-treated total RNA from PEFs and mRNA from embryos were reverse-transcribed with the High Capacity RNA-to-cDNA[™] Kit (Applied Biosystems, Foster City, CA, USA) following the manufacturer's instructions.

RT-PCR and quantitative RT-PCR

RT-PCR Synthesized cDNA from PEFs was used for PCR with 0.5 μ M of the primer pairs (Table IV.1) and the iMax 2x PCR master mix solution (iNtRON Bio Technology, Seongnam, Korea). Non-reverse transcribed RNA was used as a negative control. PCR products were run on the 1% agarose gels stained with ethidium-bromide. Each amplicon was sequenced. The reaction was carried out as

follows: one cycle of 95°C for 7 minutes; 40 cycles of 95°C for 30 seconds, 62°C for 30 seconds, and 72°C for 1 minute; one cycle of 72°C for 10 minutes.

Quantitative RT-PCR The amount of mRNA was quantified using 0.1 µM of the primer sets listed in Table IV.2 and the DyNAmo HS SYBR Green qPCR kit (Thermo Scientific, Rockford, IL) following the manufacturer's guidance. The dissociation temperature was examined and the amplicons were gel-loaded to check the reaction accuracy. *ACTB*, *RN18S*, and *YWHAG* were used as reference genes. The reaction was conducted under the following conditions: one cycle of 50°C for 5 minutes; one cycle of 95°C for 5 minutes; 40 cycles of 95°C for 15 seconds and 60°C for 1 minute.

Bisulfite sequencing

Genomic DNA was extracted from PEFs (1×10^6 cells) and zona-free parthenogenic blastocysts (n=300) using the G-spinTM Genomic DNA Extraction Kit for cells and tissues (iNtRON Bio Technology, Seongnam, Korea). Extracted gDNA was assessed using the EZ DNA Methylation-GoldTM Kit (Zymo Research, Irvine, CA) and bisulfite-treated gDNA was amplified using primer sets (Table IV.3). First-round PCR was performed with 1 cycle of 95°C for 5 minutes; 35 cycles of 95°C for 30 seconds, annealing at the temperature listed in Table IV.3 for 30 seconds; 72°C for 2 minutes; and 1 cycle of 72°C for 10 minutes. One-

microliter of the first-round PCR product was used for second-round PCR and it was performed following conditions; 1 cycle of 95°C for 7 minutes; 40 cycles of 95°C for 35 seconds, annealing temperature listed in Table IV.3 for 30 seconds, and 72°C for 1 minute; and 1 cycle of 72°C for 10 minutes. PCR product was extracted using MEGAquick-spinTM Total Fragment DNA Purification Kit (iNtRON Bio Technology, Seongnam, Korea), cloned into pGEM[®]-T Easy vector (Promega, Madison, WI), and transformed into *E.coli* (Novagen, Madison, WI). Plasmids were prepared using DNA-spinTM Plasmid DNA Purification Kit (iNtRON Bio Technology, Seongnam, Korea) and sequenced with an ABI PRISM 3730 DNA Analyzer (Applied Biosystems, Foster, CA, USA).

Table IV.1. Primer list used for RT-PCR

Gene symbol	Primer sequence (5' to 3')		Amplicon size
<i>CDX4</i>	F	CCTCGGGAAGACTGGAAC	602 bp
	R	TGCTGTATCTCAATAGGCTGAAA	
<i>CHIC1</i>	F	TGTCTCAACAAAAGAACCAGAAGA	645 bp
	R	AAAAGCAGGTAAGAGTAAAACCATT	
<i>XIST</i>	F	GCAGCTCTAAGAAGTTCCGCATT	474 bp
	R	TGTCCAGTTATCCCAAGGCATCT	
<i>LOC102165544</i>	F	CTAAGATGGCGGCGTTTG	666 bp
	R	TGGGTTTATTTCTGGGCTTT	
<i>LOC102165633</i>	F	TCAAGTTTTTACCACAAATACCA	788 bp
	R	GGCATAACAGGGACCAGAGAA	
<i>LOC100513129</i>	F	CACCAAGCTCGATCAGAAAC	688 bp
	R	GAGGAAGCACTAGGTGGAAGAA	
<i>LOC100154211</i>	F	TGCCCCATTGAAACTACCAC	500 bp
	R	CCCTCCCCCATACACACTC	
<i>SLC16A2</i>	F	ACCCCAAGCAAGAGAGGTGT	646 bp
	R	TCAGAGGGACGAACAAGAGG	
<i>LOC102166613</i>	F	TCGGTTGGATGGTGGTCT	115 bp
	R	CGATTTTCTCTCTTTCCTCCT	
<i>RLIM</i>	F	CCCACCACCGCAAACTC	567 bp
	R	CTTCCCTCAGTCTCATTTACCA	
<i>ACTB</i>	F	GTGGACATCAGGAAGGACCTCTA	131 bp
	R	ATGATCTTGATCTTCATGGTGCT	

Table IV.2. Primer pairs used in Real-time PCR		
Gene symbol	Primer sequence (5' to 3')	Amplicon size
<i>CDX4</i>	F AGCCCCTATGCGTGGATG	154 bp
	R GCTCTGATTTTCTCCGAATG	
<i>CHIC1</i>	F AATGTGAAATGGCTGCTGTG	97 bp
	R ATCTTCTGGTTCTTTTGTTGAGA	
<i>XIST</i>	F GCTCCAACCAATCTAAAAGGA	131 bp
	R ATGCCCCATCTCCACCTAA	
<i>LOC102165544</i>	F CTAAGATGGCGGCGTTTG	135 bp
	R TTGTTTTTCAGGGAATAGAGAGG	
<i>LOC102165633</i>	F GTCTGGGGTTTGTTTCCTGTG	149 bp
	R GGCTGTAGTCATCCTCTGATTTTT	
<i>LOC100513129</i>	F GCACCCCTTCACCCAGTCTT	129 bp
	R CGCAGGGCTCAATATACCTC	
<i>LOC100154211</i>	F CCACCTCCTTTTGCTGATTG	133 bp
	R TTCCCAGTCCTTTCCTCTTTT	
<i>SLC16A2</i>	F TGGTGAGGAAGACAAGGATG	183 bp
	R CCAGGAGCAGGAAGGAGATG	
<i>LOC102166613</i>	F TCGGTTGGATGGTGGTCT	115 bp
	R CGATTTTCTCTCTTTCCTCCT	
<i>RLIM</i>	F CCCACCACCGCAAACTC	159 bp
	R CGGCTCACTGCTCTCCAA	
<i>G6PD</i>	F TTCTTTGCCCGCAACTCCTA	90 bp
	R GCGTTCATGTGGCTGTTGAG	
<i>HPRT1</i>	F CATTATGCCGAGGATTTGGAA	90 bp
	R CTCTTTCATCACATCTCGAGCAA	
<i>H19</i>	F CTCAAACGACAAGAGATGGT	122 bp
	R AGTGTAGTGGCTCCAGAATG	
<i>RN18S</i>	F ACAAATCGCTCCACCAACTAAGA	90 bp
	R CGGACACGGACAGGATTGAC	
<i>ACTB</i>	F GTGGACATCAGGAAGGACCTCTA	131 bp
	R ATGATCTTGATCTTCATGGTGCT	
<i>YWHAG</i>	F CAGCCCGTGAAGATGGTG	130 bp
	R CATTGGACAGTGGCTCATTC	

Table IV.3. Primer pairs used in bisulfite-treated gDNA amplification				
Gene	Primer	Primer sequence (5' to 3')	Temp	Amplicon size
<i>CHIC1</i>	Outer-F	TTTTTGTTTTGGAGGGAAG	54°C	400bp
	Outer-R	TACTCATCACATAACCCCTTAAT		
	Inner-F	AGGGAAGGGGGTTATAATTT	54°C	360bp
	Inner-R	CCACCCTCACAAAACTCTA		
<i>XIST</i>	Outer_F	TGGTTAAATGAGGTATTTGGA	54°C	525bp
	Outer_R	CCATAAAACATAACTAAAACTAAA		
	Inner_F	TTTGTTATATTGTTGTGGAAAA	50°C	429bp
	Inner_R	CCATAAAACATAACTAAAACTAAA		
<i>LOC102165544</i>	Outer-F	TAAAAGAAATTTGGGATGGA	56°C	342bp
	Outer-R	ATTCCCAAACCTTCCTTAAC		
	Inner-F	AATTTTAGGGGTGAGAAAGG	56°C	290 bp
	Inner-R	ATTCCCAAACCTTCCTTAAC		
<i>RLIM</i>	Outer-F	TTTTTGATTATTAGTGAGGTTGAA	52°C	503 bp
	Outer-R	TTTCTTCCCCTAAAACCCCTTT		
	Inner_F	TTTGTTGTGGTTTAGTAGTAATAATTT	52°C	348 bp
	Inner_R	AAACCCCTTTATAAATCAAAA		

Statistical analysis

Statistical analysis was performed using the GraphPad Prism statistical program (GraphPad Software, San Diego, CA). Expression level comparisons between embryonic stage and somatic cells were carried out using one-way analysis of variance (ANOVA) and a Tukey test. Expression in individual fertilized blastocysts according to sex was analyzed using the Student *t*-test. Expression levels of genes in individual parthenotes, among male and female fertilized blastocysts, were analyzed by ANOVA and Dunnett's test. All data are exhibited as mean \pm standard error of the mean (S.E.M); $P < 0.05$ is considered to be significantly different.

4. Results

Identification of the XIC in pigs

To identify the XIC in pigs, orthologs of human and mouse XIC-linked genes were searched in the porcine X-chromosome. Since XIC shares a similar synteny among the eutherians, the genomic region coding *CDX4* and *RLIM*, which is located on the boundary of XIC in eutherians (Romito and Rougeulle 2011), was searched first in the pig genome. Both genes are located within 1.14 Mb on the pig X-chromosome, which is a similar length to the genomic range between human *CDX4* and *RLIM* (1.16 Mb). The orthologs of the human XIC-linked genes (*CDX4*, *CHIC1*, *XIST*, *SLC16A2*, and *RLIM*) are encoded in the 1.14 Mb porcine genomic region with the same order and transcribing direction as the human genome (Figure IV.1A and 1B). Although the orthologs of the four human XIC-linked genes (*TSIX*, *JPX*, *FTX*, and *CNBP2*) were not annotated in pigs, these four genes showed some sequence homology with the uncharacterized genes in the 1.14Mb porcine genomic region (Figure IV.1B). Similar results were obtained when mouse XIC-linked genes were compared to the porcine genome (Figure IV.2). These results showed that the composition of the genes in the porcine genomic region between *CDX4* to *RLIM* is similar to that of the XIC in other species and this region was considered to be the porcine XIC.

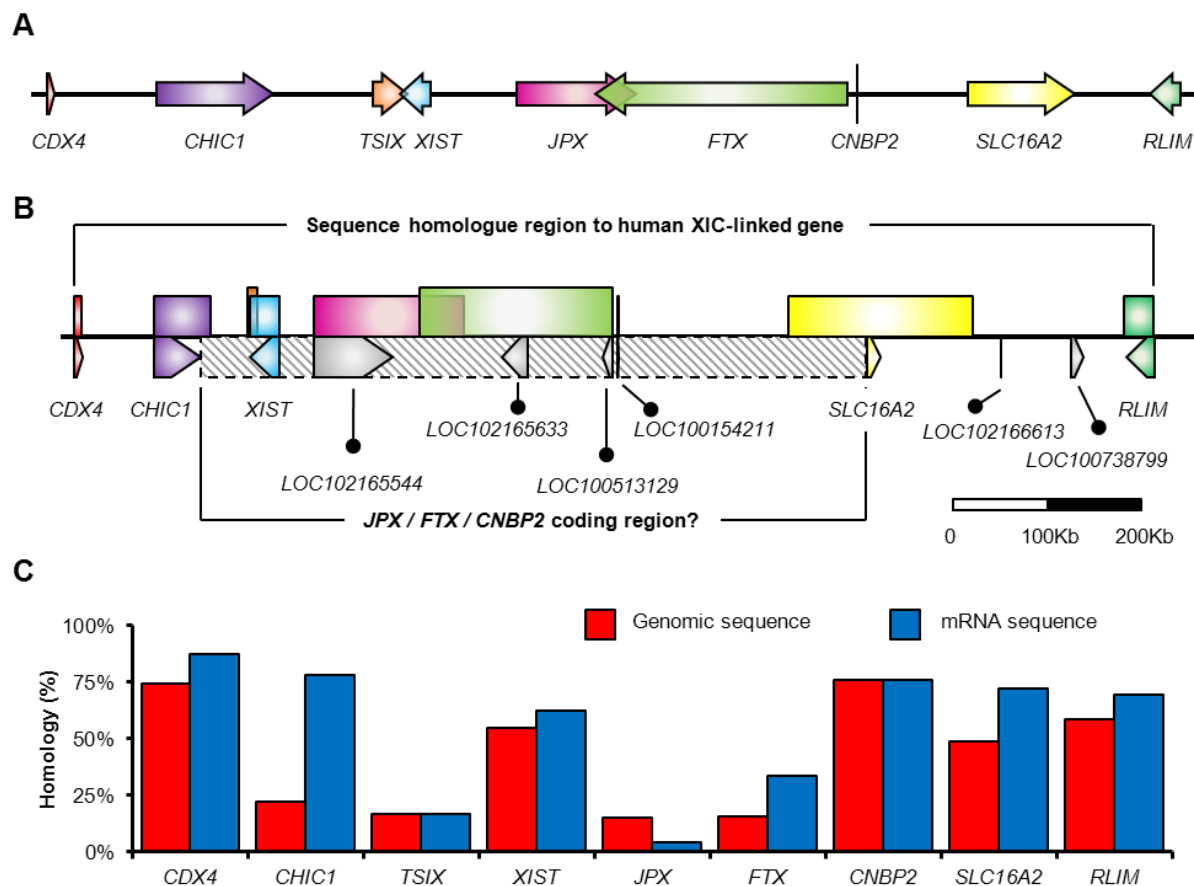


Figure IV.1. Identification of the porcine X-chromosome inactivation center (XIC) and sequence homology between human XIC-linked genes and their counterparts in pig. (A) Human XIC. Genomic region from *CDX4* to *RLIM* is represented. Arrows indicate XIC-linked genes. The diagram is scaled. (B) Candidate porcine XIC model. Arrows on bottom indicate the annotated porcine genes. Gray arrows represent uncharacterized genes. The boxes in upper panel represent the genomic range showing sequence homology with the 5'- and 3'-regions of human XIC-linked genes. Boxes with the same color as the arrows in (A) indicate the counterpart of each human XIC-linked gene in pig. The diagram is scaled. (C) Sequence homology between human XIC-linked genes and their counterpart regions in porcine XIC. Red and blue boxes indicate the homology of genomic and transcript sequences, respectively, of human XIC-linked genes to pig counterpart sequences.

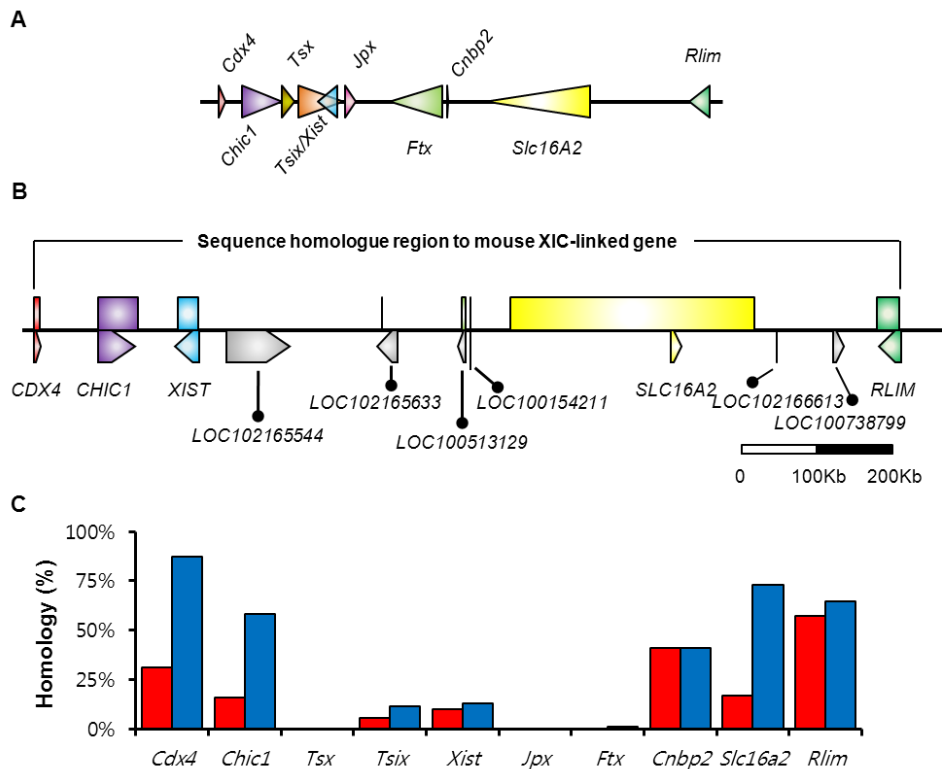


Figure IV.2. Mouse XIC and sequence comparison between porcine and mouse XIC-linked genes. (A) Mouse XIC diagram. Arrows indicate mouse genes located in the XIC. The diagram is scaled. (B) Candidate porcine XIC and the genomic region homologues to mouse XIC-linked genes. The porcine genomic regions having homology with mouse genes are represented with boxes that are the same color as their counterpart mouse genes. A detailed explanation of the diagram is stated in Figure IV.1. The diagram is scaled. (C) Sequence homology rate between mouse XIC-linked genes and their counterparts in the pig genome. Detailed information is described in Figure IV.1.

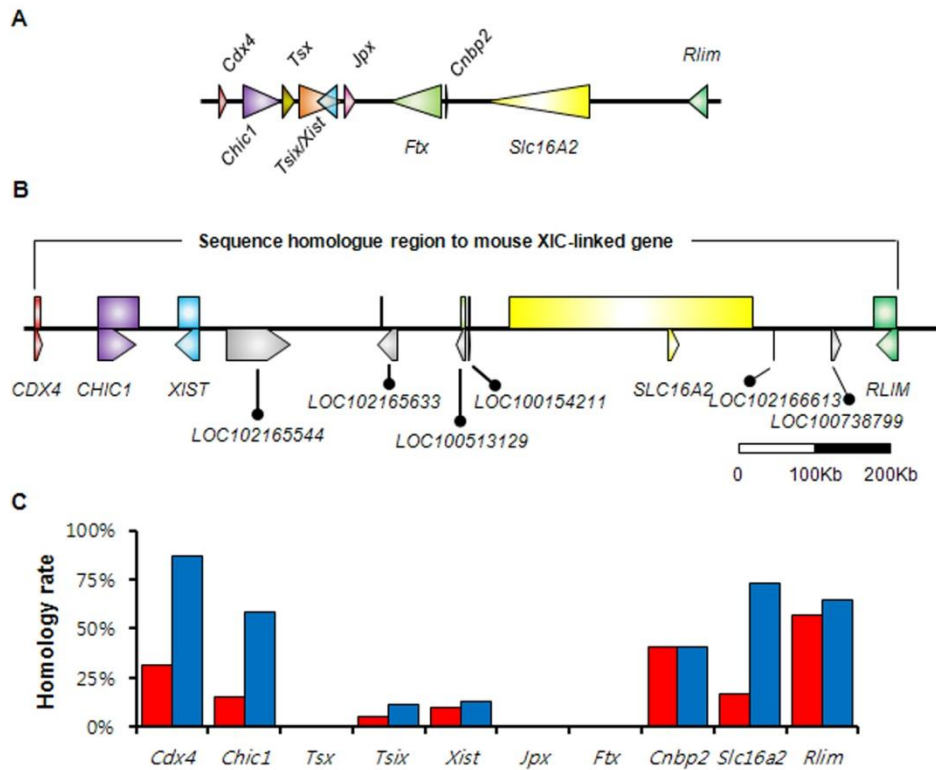


Figure IV.3. Sequence homology of the human and mouse XIC-linked gene locus with counterparts on the porcine genome. Genomic sequences of human (A) and mouse (B) XIC-linked genes were compared to their counterparts in the pig genome. Genes are represented by arrows and genes of the same color mean they are orthologous between human and mice. Bars and rectangles on the arrows represent the sequence homology of each fragment sequence and their average homology, respectively. Arrow width indicates the genomic length of the gene.

Sequences of human and mouse XIC-linked genes were compared to their counterparts in the porcine XIC using BLAST. Genomic and mRNA sequences of protein-coding genes have higher homology with their counterparts in pig rather than those of ncRNAs (Figure IV.1C and Figure IV.3). Also, each human XIC-linked gene has higher sequence homology with its counterparts in pig compared to the mouse XIC-linked genes (Figure IV.3). These results suggest that ncRNAs such as *JPX*, *FTX*, *TSIX*, and *XIST* are less conserved than protein-coding genes and that the porcine XIC is genetically closer to the human XIC rather than that of mouse.

Expression of XIC-linked genes in embryonic fibroblasts

In my previous study, it was found that one of the X-chromosome is accomplished and the inactive status is maintained in porcine embryonic fibroblasts (PEFs) because one *XIST* promoter allele is methylated in female PEFs (Hwang *et al.* 2013a). Therefore, the expression levels of XIC-linked genes were analyzed in male and female PEFs to confirm the compensated expression and presence of XCI-escaping genes like *XIST*. The expression levels of 10 porcine XIC-linked genes were examined and two genes, *LOC10738799* and *LOC100738472*, were excluded because the mRNA sequences of these two uncharacterized genes are identical to porcine *SLC16A2* and *RLIM*, respectively. Of the examined XIC-linked genes, nine genes except *CDX4* were expressed in PEFs and *XIST* were the only detected genes in female PEFs (Figure IV.4A). Sexual dimorphic expression was

not observed in most XIC-linked genes but *XIST* and *LOC102165544* were expressed preferentially in females (Figure IV.4B). Interestingly, *LOC102165633*, which has some sequence homology with human *JPX* (Figure IV.1 and Figure IV.3), was expressed 2-fold higher in female PEFs compared to male PEFs (Figure IV.4B). These results indicate that the dosages of overall XIC-linked genes are compensated between male and female PEFs, and *XIST* and *LOC102165544* are genes that escape XCI in pigs.

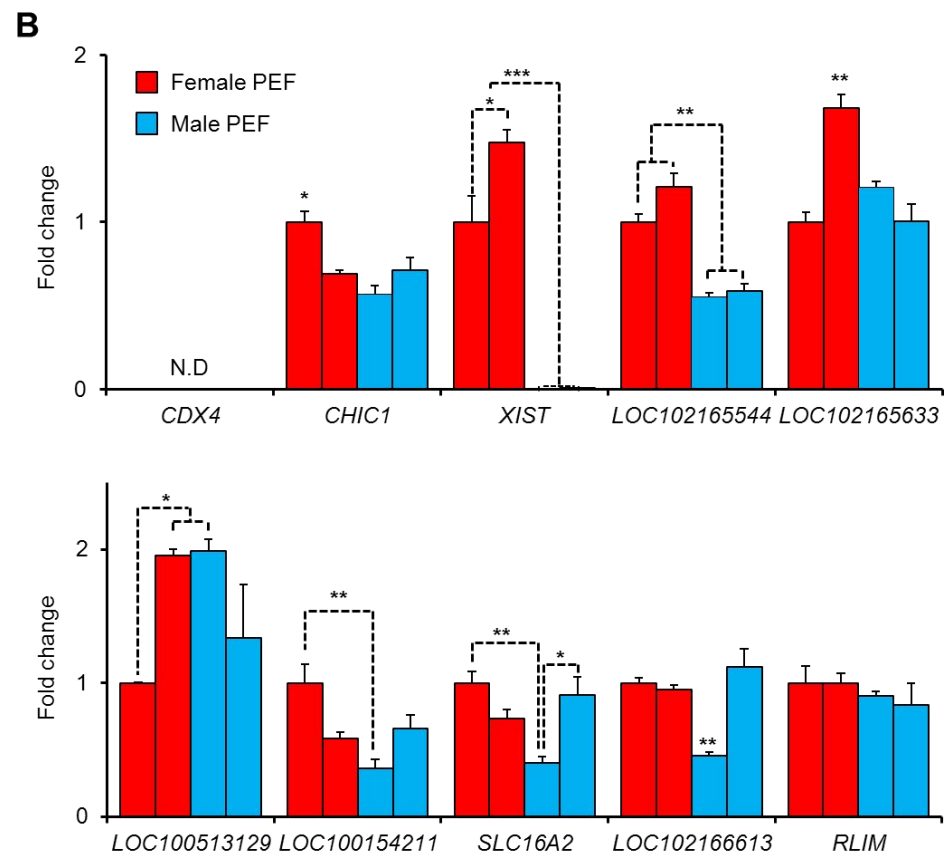
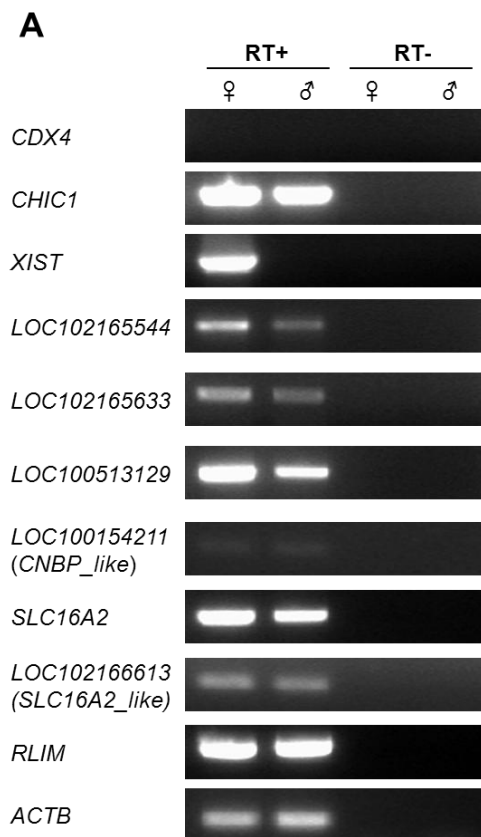


Figure IV.4. XIC-linked gene expression in porcine embryonic fibroblasts (PEFs). Expression of XIC-linked genes (A) and their quantitative comparison (B) in male and female PEFs were performed. Non-reverse transcribed RNA was used as a negative control. Ct-value of *ACTB* was used to normalize each gene expression level. The expression amount of XIC-linked genes in one female PEF (first column in panel B) set to 1-fold.

Expression of XIC-linked genes in parthenogenic embryo development

XCI occurs by accumulation of *XIST* on one of the X-chromosomes in early eutherian embryo development (Penny *et al.* 1996; Wutz and Jaenisch 2000; Okamoto *et al.* 2004; Okamoto *et al.* 2011). The genes located on the XIC, such as *Rlim*, *Jpx*, and *Ftx*, have been suggested to be regulators of XCI and *Xist* expression in mice (Jonkers *et al.* 2009; Tian *et al.* 2010; Chureau *et al.* 2011). These studies suggest the importance of the XIC-linked genes for regulating XCI and thus, the expression level changes of XIC-linked genes were examined in preimplantation embryos. To assess female embryos in each stage, parthenogenic embryos were used in this study. Among the 10 XIC-linked genes examined in PEFs, *CHIC1*, *XIST*, *LOC102165544*, and *RLIM*, were consistently expressed in embryonic stages (data not shown). The expression levels of *XIST* and *LOC102165544* were increased (*XIST*) or maintained (*LOC102165544*), but the expression of *CHIC1* and *RLIM* were decreased in blastocysts compared to the morulae (Figure IV.5). As the increase and accumulation of *XIST* is essential for initiating XCI (Penny *et al.* 1996), the expression levels of the four genes were compared between individual male and female blastocysts to confirm whether the dosage of the each gene was compensated in blastocyst stage.

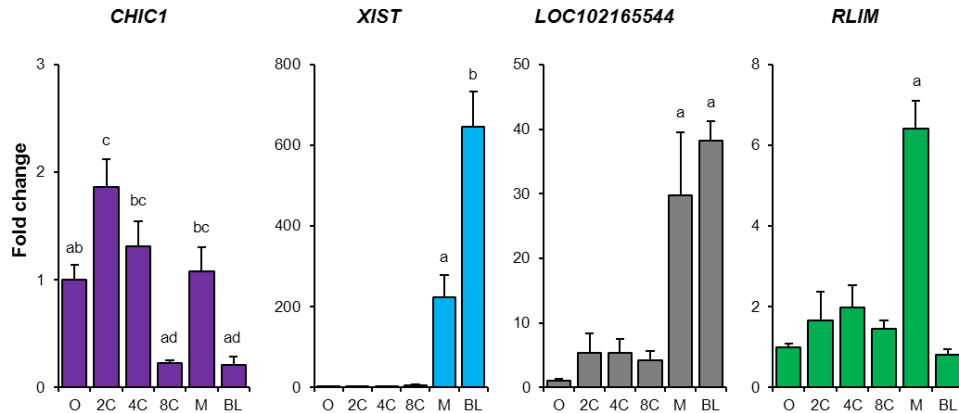


Figure IV.5 Porcine XIC-linked gene expression at the embryonic stage.

Expression level changes of four XIC-linked genes were measured in each stage of parthenogenic preimplantation embryos. Pooled oocytes (O, n=40), 2-cells (2C, n=20), 4-cells (4C, n=20), 8-cells (8C, n=20), morulae (M, n=10), and blastocysts (n=5) were used. Ct-value of *YWHAG* was used for normalize each gene expression in embryos and oocytes. Expression levels of genes in oocytes were set to 1-fold. Different characters indicate significant difference ($P<0.05$). The reaction was replicated four times.

Expression level comparison of XIC-linked genes in male and female blastocysts

Expression levels of the four XIC-linked genes were compared in male and female blastocysts to confirm that the dosage of each gene was compensated in blastocysts (Figure IV.6). The sex of the *in vitro* fertilized blastocysts was determined by *XIST* expression level (Park *et al.* 2011) prior to analyzing the expression of the genes. Also, the expression of two X-linked genes, *HPRT1* and *G6PD*, which were reported to be preferentially expressed in female blastocysts (Park *et al.* 2012), were examined in sexed blastocysts to further confirm that embryo sexing was performed properly. Sexed female blastocysts in this study had 1.6-fold higher expression of *HPRT1* and *G6PD* compared to male blastocysts, as previously reported. This result shows that the two groups of blastocysts classified by *XIST* expression level represent each sex well. *LOC102165544*, a gene that did not have decreased expression levels during morula to blastocyst development, was highly expressed in female blastocysts. However, the expression levels of *CHIC1* and *RLIM*, which showed decreased expression at the blastocyst stage compared to morula stage, were not significantly different between male and female blastocysts. This result reveals the possibility that dosage compensation is achieved in only a few XIC-linked genes in porcine blastocysts.

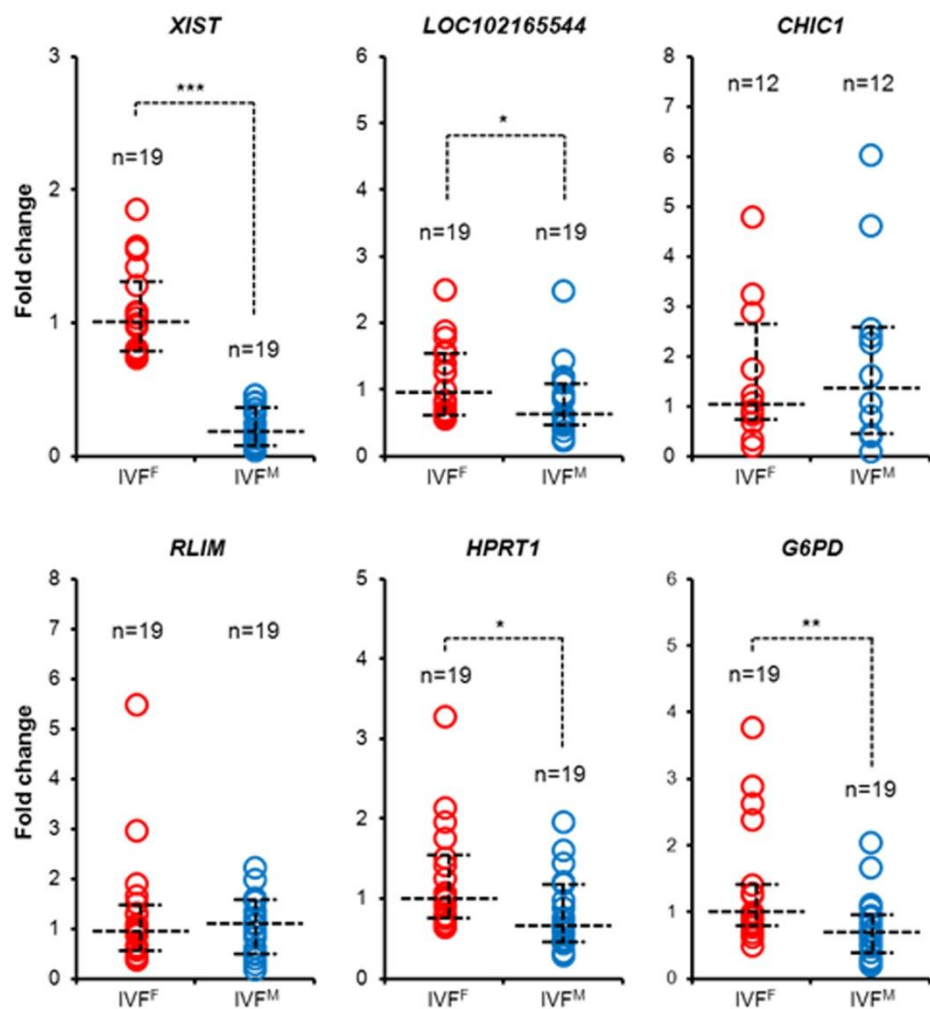


Figure IV. 6. Quantitative comparison of XIC-linked genes expression in male and female blastocysts. The expression amount of each XIC-linked gene was compared between male and female blastocysts produced by *in vitro* fertilization. *HPRT1* and *G6PD*, which were previously reported to be expressed preferentially in female porcine blastocyst (Park *et al.* 2012), were used to confirm sexing of fertilized blastocysts. Each red and blue circle indicates the expression level of individual female and male blastocysts, respectively. Box-plots in each column represent the quartile distribution of the gene expression levels of each blastocyst. *ACTB* and *RN18S* were used as reference genes, and the median of normalized expression levels among female blastocysts was determined to 1-fold. Asterisks indicate significant difference between male and female blastocysts (*, $P<0.05$; **, $P<0.01$; and ***, $P<0.001$).

CpG sites methylation status of XIC-linked genes in blastocysts.

The methylation status of the CpG sites of the four genes was examined in female porcine blastocysts to confirm that the compensated dosages of *CHIC1* and *RLIM* are induced by inactivation of one allele. The CpG sites of each XIC-linked genes were searched (Figure IV.7) and examined their methylation patterns in male and female PEFs and parthenogenic blastocysts. CpG sites of three X-linked genes, *CHIC1*, *RLIM*, and *XIST*, showed sex dependent methylation patterns in PEFs, but *LOC102165544* was hypomethylated in both sexes (Figure IV.8A and 8B). This result supports that *LOC102165544* would be an XCI-escaping gene resulted from transcriptional analysis (Figure IV.4B). Contrary to female PEFs, each CpG site was generally hypomethylated in parthenogenic blastocysts except *CHIC1* (Figure IV.8C). The CpG sites of *XIST* and *LOC102165544*, which were differentially expressed between male and female blastocysts, were fully demethylated in parthenogenic blastocysts. The bi-allelic active promoter of *XIST* indicates that XCI is not accomplished in female porcine blastocysts. The CpG sites of *CHIC1* and *RLIM*, which showed quantitatively matched expression between male and female blastocysts, showed differential methylation patterns in female blastocysts. The CpG sites of *CHIC1* were about half methylated but those of *RLIM* were demethylated in female blastocysts. These results mean that dosage compensation of X-linked genes wouldn't be completed throughout the entire chromosome although a few genes are considered to initiate inactivation of their one allele in female porcine blastocyst.

***XIST* expression is sex dependent in blastocysts**

Analysis of XIC-linked genes expression and methylation patterns supports the possibility that dosage compensation of a few XIC-linked genes is achieved in porcine blastocysts. However, it is important to confirm whether the expression of *XIST* is dependent on the parental origin of the allele because *XIST* accumulation is essential for inducing XCI and imprinting XCI wasn't observed in rabbit and human embryos. Therefore, the expression of *XIST* and other X-linked genes were compared between parthenogenic and *in vitro* fertilized blastocysts to confirm the effect of the origin of the allele in the expression of each XIC-linked gene (Figure IV.9). One study showed that *XIST* is preferentially expressed on the maternal X (Xm) in porcine blastocysts. (Park *et al.* 2011). However, this result is unclear because the analysis was conducted using pooled blastocysts without distinguishing the sex of each blastocyst. Comparison of expression levels of *XIST* using sexed individual blastocysts showed that *XIST* expression levels are similar between parthenogenic and *in vitro* fertilized female blastocysts, contrary to the previous study. This means that *XIST* expression is not dependent of the origin of the allele, but is sex-dependent in porcine blastocysts. Unexpectedly, *LOC102165544* expression was significantly higher in *in vitro* fertilized female blastocysts having the Xp. This would mean *LOC102165544* is preferentially expressed in Xp preferentially in porcine blastocysts. Other XIC-linked genes, *CHIC1* and *RLIM*, also showed that expression is not dependent on the origin of allele.

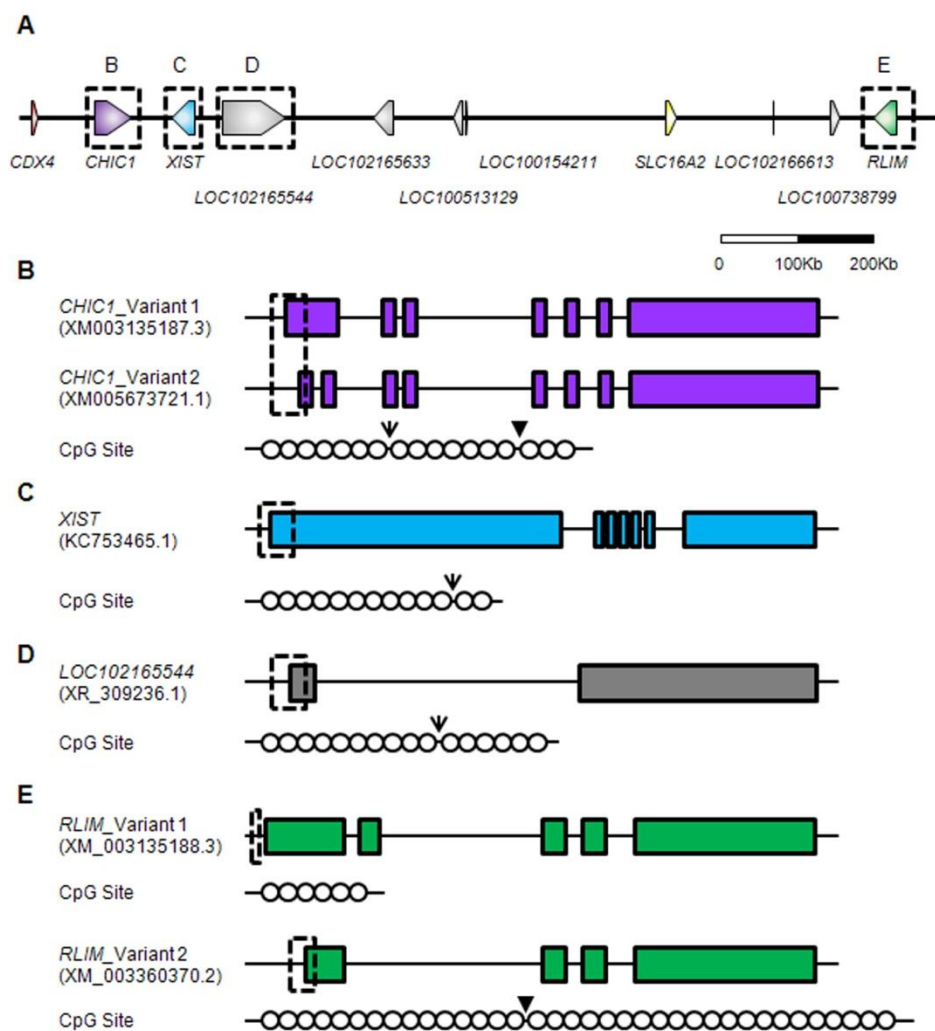


Figure IV.7. Diagram of CpG sites of genes in the porcine XIC. (A) Diagram of the porcine XIC. Detailed information is described in Figure IV.1. (B-E) Gene structure and identified CpG sites of each gene. The structure of *CHIC1* (B), *XIST* (C), *LOC10216554* (D), and *RLIM* (E) is depicted. A rectangle indicates the exon of the gene and circles represent CG dinucleotides. Arrows and arrowheads indicate the first and second transcription starting sites, respectively. The CpG sites of the *RLIM* variant 1 was considered not to be a proper differentially methylated region because the region was identically methylated in both sexes of PEFs (data not shown).

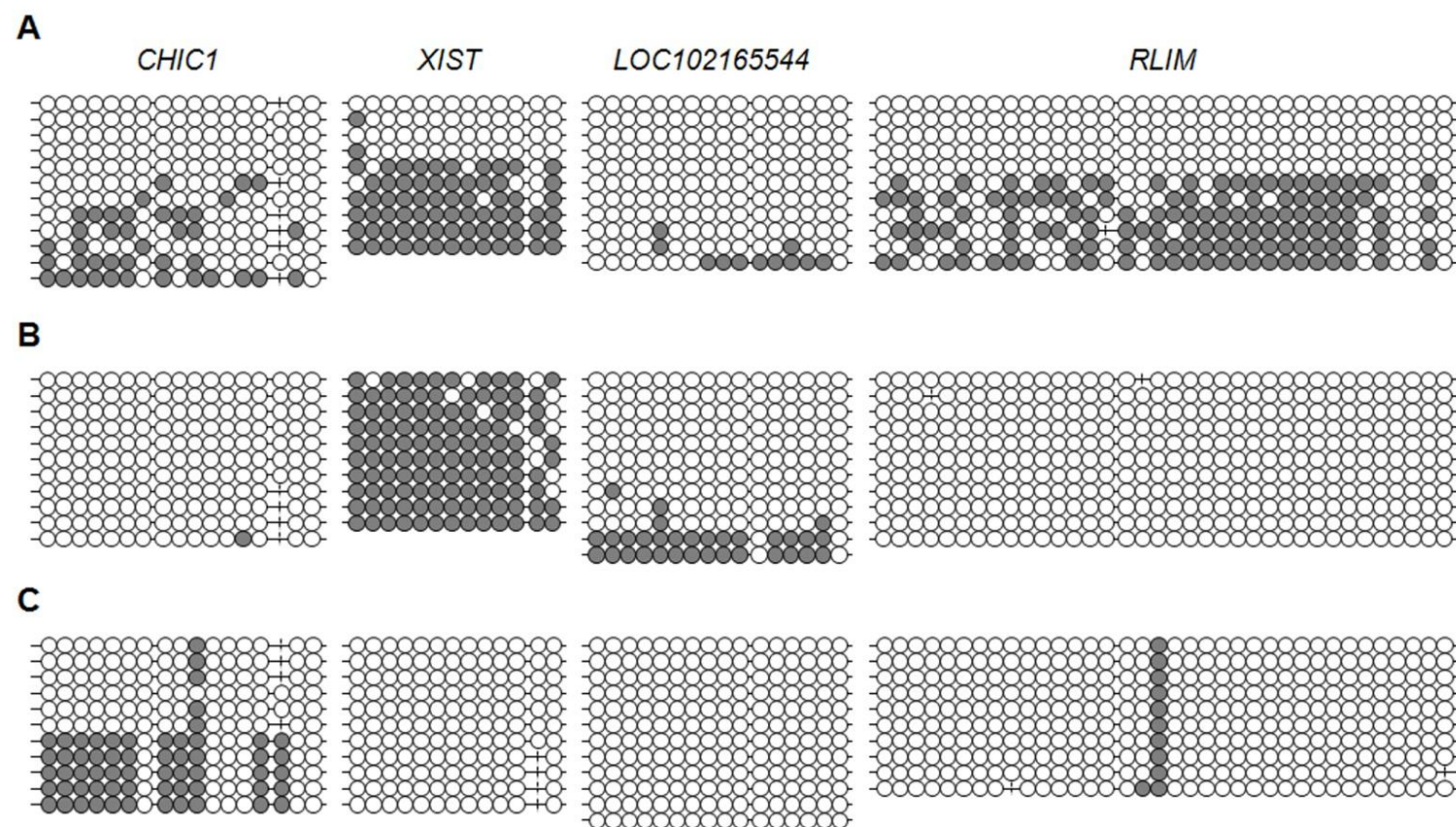


Figure IV.8. Methylation status of XIC-linked genes CpG sites in porcine embryonic fibroblasts (PEFs) and parthenogenic blastocysts. Methylation patterns of CpG sites in each XIC-linked gene were examined in female (A) and male (B) PEFs, and parthenogenic blastocysts (C). CpG sites of the XIC-linked genes were identified near the transcription starting site of the genes (Figure IV.7). Each open and closed circle indicates demethylated and methylated cytosine-guanine (CG) dinucleotides, respectively. The diagram is not scaled.

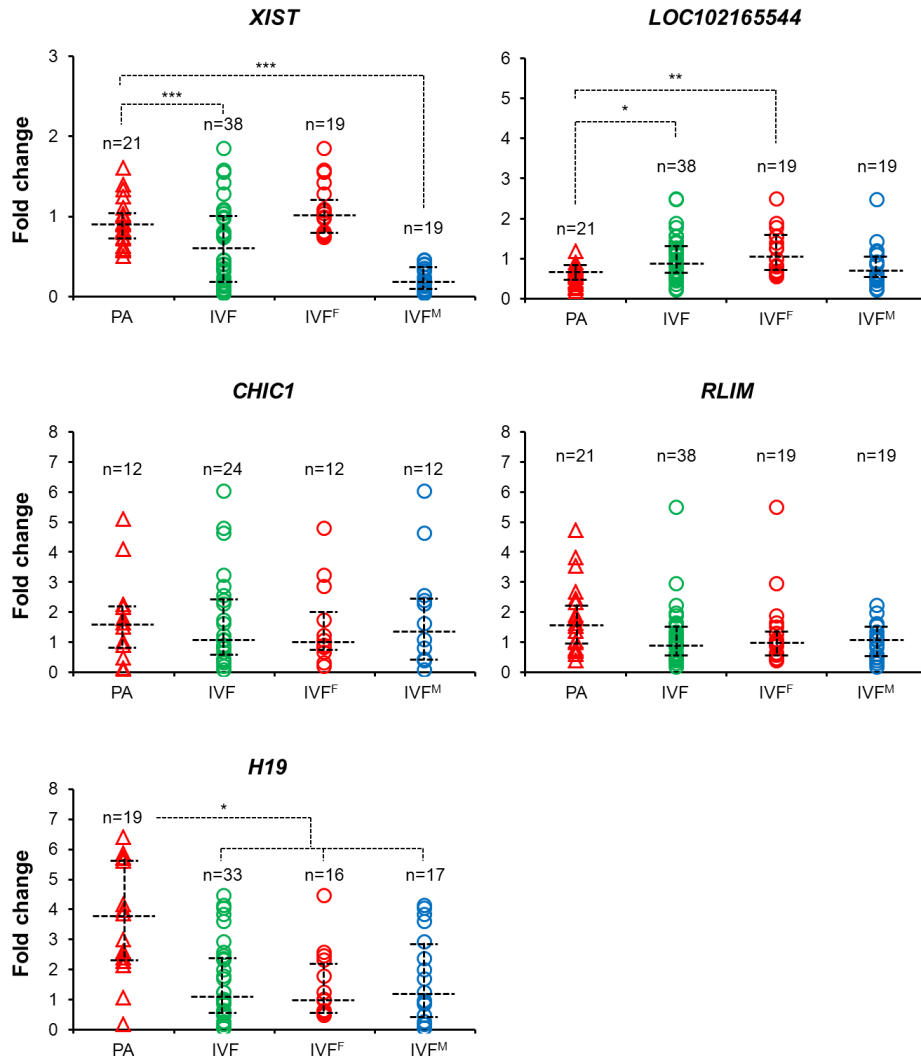


Figure IV.9. Quantitative comparison of expression levels of XIC-linked genes in porcine parthenogenic and fertilized blastocysts. Expression levels of XIC-linked genes in individual parthenogenic blastocysts (red triangles) were compared according to sex of *in vitro* fertilized blastocysts. *H19* was used to represent paternal imprinting gene in pigs (Park *et al.* 2009; Park *et al.* 2011). Green circles indicate both male and female *in vitro* fertilized blastocyst. Other details are described in the legend of Figure IV.6.

5. Discussion

XCI is an essential process for compensating the dosages of X-linked genes between the male and female eutherians. The event occurs during the early phase of embryo development with various strategies among eutherians (Escamilla-Del-Arenal *et al.* 2011; Okamoto *et al.* 2011). Despite diverse and species-specific manners for achieving dosage compensation of X-linked genes, this epigenetic event has only been confirmed in a few species, and little is known regarding the process in pigs.

XCI is induced mainly by *Xist*, and various factors are suggested to regulate *Xist* expression in mice. Among the factors, XIC-linked genes such as *Jpx*, *Ftx*, and *Rlim* have been reported to induce and regulate *Xist* expression (Jonkers *et al.* 2009; Tian *et al.* 2010; Chureau *et al.* 2011). Compared to an study that defined XIC more than thirty years ago (Rastan 1983) prior to identifying *XIST/Xist* (Brockdorff *et al.* 1992; Brown *et al.* 1992), recent reports focusing on XIC-linked genes demonstrate that not only *Xist*, but other XIC-linked genes are involved in XCI. Although the importance of the XIC-linked genes has been highlighted in various reports, the porcine XIC has not been identified. Thus in this study, XCI status was examined by focusing on the XIC-linked genes in porcine blastocysts.

Conserved XIC consists of less conserved non-coding RNA genes in pigs.

Comparative sequence analysis revealed that the XIC is generally conserved because of gene synteny and order despite the differential genomic length of the region between human and mice (Chureau *et al.* 2002). Among the XIC-linked genes, orthologs of protein-coding genes, *CDX4*, *CHIC1*, and *SLC16A2*, have been observed in non-eutherian vertebrates and they are located in proximity to each other (Duret *et al.* 2006). However, the genes comprising the genomic region between *CHIC1* and *SLC16A2* are different between eutherian and non-eutherian vertebrates. The internal region is composed of five protein-coding genes in non-eutherian vertebrates, but their orthologs were not defined in eutherians. This counterpart region in eutherians contains ncRNAs instead (Duret *et al.* 2006). These differences suggest that an evolutionary break-point of XIC-linked genes in the genomic region between *CHIC1* and *SLC16A2* emerged during adaptive radiation of mammals by disruption of the protein-coding genes (Hore *et al.* 2007; Romito and Rougeulle 2011). The disruption of the ancestral protein genes by integration of a mobile element, or truncation of a coding region resulted in pseudogenization, which lead to the emergence of eutherian-specific ncRNAs like *XIST* (Elisaphenko *et al.* 2008). A previous report confirmed that mouse *Tsx* is pseudogenized and the *Ppnx* ortholog was not detected in humans (Chureau *et al.* 2002). These studies indicate that ncRNAs in XIC emerged during an evolutionary short period, which resulted in less sequence homology of the ncRNAs among the eutherian species, even though the orthologs of XIC-linked genes are present in

various species including non-eutherian vertebrates.

In this study, the sequence comparison between human XIC-linked genes and their pig counterparts showed that ncRNAs are less conserved compared to protein coding genes (Figure IV.1). Low sequence similarity with mouse XIC-linked genes and an absence of the *Tsx* homologue region in the porcine XIC suggests that the porcine XIC is close evolutionary to that of human XIC rather than the mouse XIC (Figure IV.1 and Figure IV.2). Sequence comparison between human and pig showed that *XIST* is the most sequence-conserved gene among the XIC-linked ncRNAs. This would mean that the role of *XIST* is most conserved compared to the other ncRNAs.

Human *JPX* and *FTX*, whose orthologs are a positive regulator for *Xist* expression in mice (Tian *et al.* 2010; Chureau *et al.* 2011), showed little sequence homology with their counterparts in pigs. Uncharacterized ncRNA genes coded between the *XIST* and *SLC16A2* in pig are considered orthologs for *JPX* (*LOC102165544*) and *FTX* (*LOC102165633* or *LOC100513129*) because of the coding region and transcription strand. However, sequence comparison showed that these two genes are the least conserved among the XIC-linked genes in pig and this result suggests that the uncharacterized genes may not be their orthologs in pigs. However, because sequence homology doesn't directly reflect functional conservation (Pang *et al.* 2006), additional studies regarding the role of the uncharacterized ncRNAs in XCI are required.

The presence of porcine *TSIX*, whose ortholog is known to be a negative regulator for *Xist* in mice, (Lee *et al.* 1999) is unclear because only the region antisense to the 3' region of porcine *XIST* showed partial sequence homology with the 5' region of human *TSIX*, which overlapped to the 3' region of human *XIST* (Figure IV.3A). Although the human *TSIX* was identified more than 10 years ago (Migeon *et al.* 2001), it is still unclear if human *TSIX* is a negative regulator for *XIST*, as shown with mouse *Tsix*. This suggests that the ortholog is a less conserved gene compared to the *XIST* ortholog. Therefore, considering the results and the uncertain function of human *TSIX*, porcine *TSIX* might not be present or may be expressed in another genomic region. Thus, identifying *TSIX* orthologs in pig is still needed. Common features among the species, such as conserved synteny and sequences of protein-coding genes and less conserved ncRNAs, show that the examined genomic region is the porcine XIC.

XCI is initiated but not accomplished across the entire chromosome in porcine blastocysts.

A recent report demonstrated that the time-window for XCI initiation and completion varies among species (Okamoto *et al.* 2011). The report highlighted that rabbits and humans have bi-allelic expression of *XIST* in early blastocysts and these two species do not have imprinting XCI during preimplantation embryo development. Interestingly, bi-allelic *XIST* accumulation is mainly observed in

human blastocysts, and a few X-linked genes were active in both alleles. This means that the dosages of some X-linked genes, at least the examined three genes in the report, are not compensated between male and female human blastocysts. However, another study suggested that XCI is initiated in human blastocysts with the evidence that *CHIC1* in a *XIST*-coated X-chromosome was silenced in human blastocysts (van den Berg *et al.* 2009). Although the *XIST* expression patterns in human blastocysts were different between the two studies, these reports showed that the allele inactivation timing is different among the X-linked genes. This feature is also observed in bovine blastocysts (Bermejo-Alvarez *et al.* 2010). Approximately 20% of X-linked genes are identically expressed between male and female bovine blastocysts and the remaining 80% are expressed differently between the two sexes. This differential expression status of X-linked genes caused by both or a single active allele of the genes in human and bovine female blastocysts suggests that an extended period for completing XCI is needed compared to the process in mice. Indeed, elongated bovine embryos showed a higher proportion of compensated X-linked gene expression compared to spherical blastocysts but expression levels of some of X-linked genes were still not balanced between male and female (Bermejo-Alvarez *et al.* 2011). This report supports the idea that XCI completion in bovine embryo development requires a longer time compared to mice.

Results in this study show that the dosages of two X-linked genes are compensated between male and female blastocysts (*CHIC1* and *RLIM*) (Figure

IV.6). Also, the methylation status showed that one allele of *CHIC1* is inactivated in female porcine blastocysts (Figure IV.8). These results support the possibility that dosage compensation of subtle XIC-linked genes is initiated and achieved in porcine blastocysts. This is also supported by the result that the highest expression of *RLIM* is in morulae (Figure IV.5) because the gene is reported to induce XCI in a dose-dependent manner and down regulated after XCI initiation by inactivation of its one allele on inactive X-chromosome in mice (Jonkers *et al.* 2009). However, it does not mean that the inactivation of all X-linked genes in future inactive X-chromosome is achieved in this embryonic stage because sexual-dimorphic expression of other XIC- (*XIST* and *LOC102165544*) and X-linked genes (*HPRT1* and *G6PD*) was observed in this study (Figure IV.6) and in previous studies (Park *et al.* 2012). Demethylated CpG sites of *XIST* also support the idea that porcine XCI is not completed at the blastocyst stage (Figure IV.8). One report confirmed that accumulation of H3K27ME3 is detected first in elongated porcine embryo (Gao *et al.* 2011). Considering that the epigenetic chromatin modification means XCI completion, the report further supports the idea that XCI is not completed in porcine blastocysts. Therefore, it is considered that XCI is not accomplished in porcine blastocysts even though the dosages of a few genes are compensated between males and females (Figure IV.10). Contrary to previous study (Park *et al.* 2011), *XIST* is expressed sex dependently rather than parental origin of X-chromosome (Figure IV.9). This difference in expression would be arisen by pooled blastocyst without distinguishing sexes in the previous study. This result

indicates that *XIST* may not be the imprinting gene in pigs. As this expression would be related to the random or imprinted XCI in pigs, additional studies are needed to clarify the presence of imprinting XCI. These results support the idea that dosage compensation of a few X-linked genes is achieved but chromosome-wide inactivation of X-linked gene would not be accomplished in porcine blastocyst, and the time-window and process of XCI in pig is similar to that of human rather than mouse.

Uncharacterized gene, *LOC102165544*, as the possible *Jpx* ortholog in pigs.

Tian et al. revealed that an ncRNA, *Jpx*, is the trans-activator for inducing *Xist* expression and XCI in differentiation of mouse embryonic stem cells (ES cells) (Tian *et al.* 2010). *Jpx* lies about 10 kb upstream from *Xist* and is transcribed opposite to *Xist*. Expression of this XCI-escaping ncRNA is increased during differentiation of mouse ES cells, and its deletion induces failure of XCI in female ES cells. One recent study also reported that the dosage of *Jpx* is essential for inducing XCI by titrating CTCF and activating one of the *Xist* alleles during mouse ES cell differentiation (Sun *et al.* 2013).

In this study, the genomic location and expression pattern of *LOC102165544* were similar to that of *Jpx*: a genomic distance of about 40 Kb from *XIST* on the opposite transcribing strand (Figure IV.1B). There is approximately 2-fold higher

expression in female blastocysts and PEFs (1.6- and 1.94-fold, respectively) compared to males (Figure IV.4 and Figure IV.6) and the methylation pattern in PEFs (Figure IV.8) is evidence that this gene escapes XCI. Also, the increase of expression following *XIST* expression at the embryonic stage (Figure IV.5) is similar to *Jpx* expression in differentiating mouse ES cells. These similarities between *Jpx* and *LOC102165544* support the possibility that this uncharacterized gene is the *Jpx* ortholog in pigs. However, it is hard to exclude another possibility that this ncRNA is not the *Jpx* ortholog because of the low sequence homology and differential gene structure to *Jpx* (Figure IV.1C). Also, the Xp-preferential expression of *LOC102165544* is a different expression pattern compared to *Jpx* (Figure IV.9). Therefore, further examination and characterization of the function of this gene in porcine XCI is needed.

In this study, the porcine XIC was identified and dosage compensation of XIC-linked genes was examined in porcine preimplantation embryos. The results demonstrate that the porcine XIC shares genomic character with other species. Also, the dosage compensation of some XIC-linked genes would be achieved, although chromosome-wide XCI was not considered to be accomplished in blastocysts in this study. However, there are several limitations of this study. At first, only *in vitro* embryos were utilized and examined selected genes in here. Also, gene expression analysis was carried out using whole blastocysts without distinguishing between

differential lineage cells such as ICM and trophectodermal cells. Therefore, cytological analysis using RNA FISH and examination of global gene expression using *in vivo* embryos are needed to confirm the exact time-window for XCI initiation and mechanism in pigs.

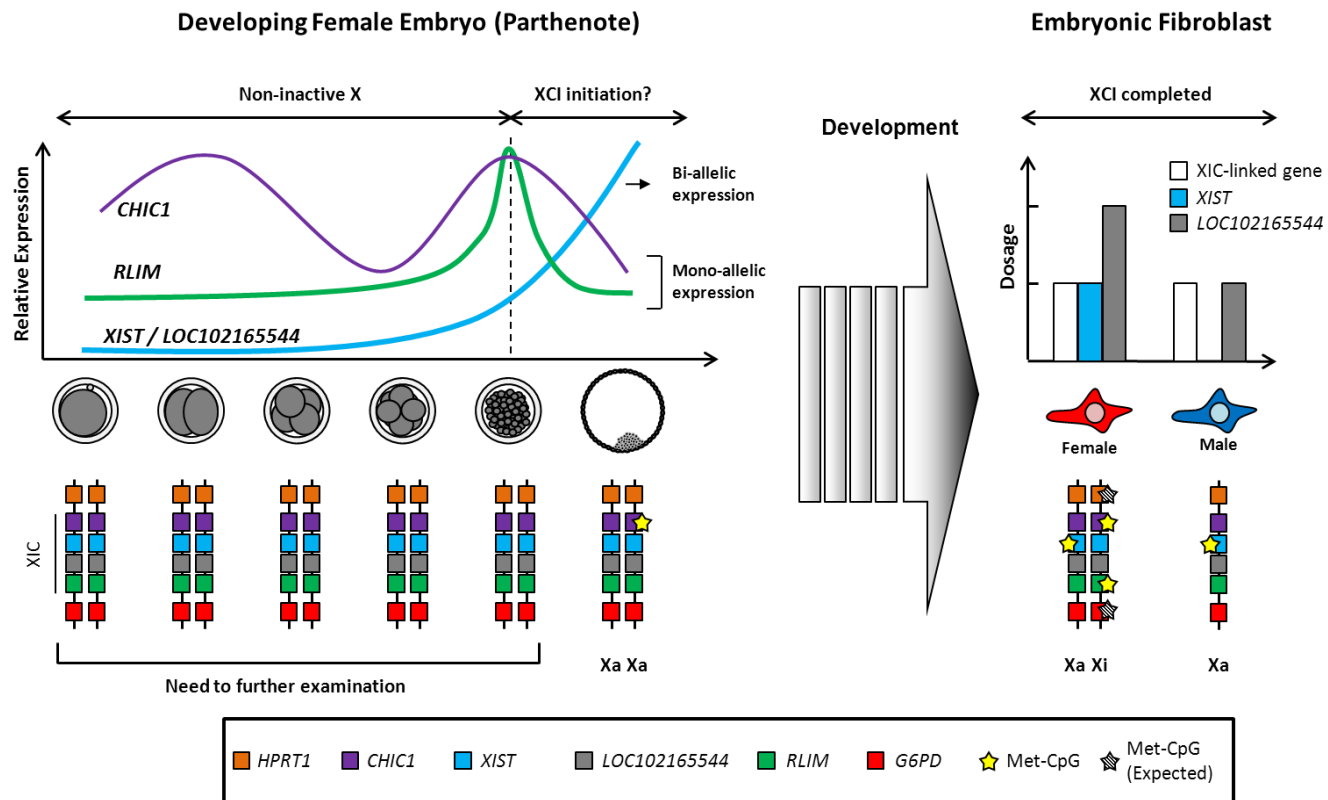


Figure IV.10. Epigenetic dynamics of XCI status and XIC-linked gene expression in porcine developing embryos and fibroblasts

CHAPTER V

Identification and differential expression patterns

of porcine *OCT4* variants

1. Abstract

OCT4 encoded by *POU5F1* has a crucial role of maintaining pluripotency in embryonic stem cells during early embryonic development and several *OCT4* variants have been identified in mouse and human studies. The objective of the present study was to identify different variants of *OCT4* and analyze their expression patterns in preimplantation porcine embryos and in various tissues. Here, it was confirmed that porcine *POU5F1* transcribes its three variants, namely *OCT4A*, *OCT4B*, and *OCT4B1*. The *OCT4B* transcript consists of exons identical to the major form of the *OCT4* variant, *OCT4A*, with a differential N-terminal domain coding exon. The structure of *OCT4B1* mRNA was same as that of *OCT4B* mRNA, but harbored a cryptic exon. Based on these findings, the transcription levels was investigated and found that *OCT4B* and *OCT4B1* made up approximately 20% among the variants in embryonic stage and this indicates *OCT4A* mRNA is dominantly expressed during preimplantation embryos development. Also, *OCT4B* mRNA was detected in all tissues examined, while *OCT4A* and *OCT4B1* were detected only in testis but not in other tissues examined. *OCT4B1* showed inversely correlated expression with *SOX2* and *NANOG* expression. *OCT4A* protein was specifically localized in the nuclei, whereas *OCT4B* was mainly localized in the cytoplasm of the porcine embryos at the blastocyst stage. The findings of this study reveal that the porcine *OCT4* gene can

potentially encode three variants (*OCT4A*, *OCT4B*, and *OCT4B1*), and they are differentially expressed and would have dissimilar roles each other in preimplantation embryos and various adult tissues.

Key words: Gene expression, Embryo development, *OCT4* variant

2. Introduction

The transcription factor *OCT4* (also named *POU5F1*) belongs to a protein family encoding the octamer-binding POU (Pit-Oct-Unc) domain (Scholer *et al.* 1990; Takeda *et al.* 1992). The *Oct4* gene has been shown to be expressed in the early developmental embryonic stage, embryonic carcinoma (EC) cell, embryonic stem (ES) cell, and germ cell lineage (Okamoto *et al.* 1990; Rosner *et al.* 1990; Pesce *et al.* 1998). *Oct4A*, which is the most well known form among the *Oct4* variants because it is a representative marker for pluripotency, has been suggested to be a gatekeeper for maintaining pluripotency in embryonic stem (ES) cells and is essential for inner cell mass (ICM) formation during embryo development in mouse (Nichols *et al.* 1998; Niwa *et al.* 2000). Deletion of *Oct4* induces failure of pluripotent ICM formation in the blastocyst (Nichols *et al.* 1998) and down- or up-regulation of *Oct4* makes ES cells differentiate into trophectodermal cells or mesodermal cells, respectively (Niwa *et al.* 2000). The *OCT4* gene has also been suggested as one of the essential transcription factors for reprogramming somatic cells and inducing those to have pluripotency (Takahashi and Yamanaka 2006; Takahashi *et al.* 2007). These studies have clearly demonstrated the importance of the *OCT4* gene, especially the A form, in embryo development and ES cell pluripotency.

Human and mouse *OCT4* variants, *OCT4B* and *OCT4B1*, have been identified (Takeda *et al.* 1992; Atlasi *et al.* 2008; Mizuno and Kosaka 2008) and their expressions have been examined in preimplantation embryos, ES cells, embryonic carcinoma (EC) cells, somatic cells, and various cancer cells (Cauffman *et al.* 2006; Lee *et al.* 2006b; Atlasi *et al.* 2008; Mizuno and Kosaka 2008; Guo *et al.* 2012). The *OCT4B* transcript shares identical exons coding the POU DNA-binding and C-terminal transactivation domains (CTDs) of *OCT4A* but the variants have different exon-coding N-terminal transactivation domains (NTDs) (Takeda *et al.* 1992; Mizuno and Kosaka 2008). And the *OCT4B1*, which is composed same exons with *OCT4B*, contains an additional cryptic exon which is an intron region in *OCT4A* and *OCT4B* transcript (Atlasi *et al.* 2008; Wang and Dai 2010).

The functions of the variants of *OCT4* are still poorly understood. The relationship between pluripotency and the *B* and *B1* forms of *OCT4* has been reported in only a few studies (Lee *et al.* 2006b; Atlasi *et al.* 2008; Guo *et al.* 2012; Mirzaei *et al.* 2014). These studies showed that *OCT4B* does not appear to be related with pluripotency (Lee *et al.* 2006b; Guo *et al.* 2012). However, one study originally defining *OCT4B1* in human, suggested that the *OCT4B1* variant could have a relationship with pluripotency by comparing expression levels among EC, ES, and somatic cells. These studies suggested that the novel *OCT4* variants would have distinctive functions compared to *OCT4A*. However, little information exists so far on the on *OCT4* expression of the different variants in pig species. Therefore, the present study established undefined variants, *OCT4B* and *OCT4B1*, and

determined the expression patterns of them in various somatic tissues and embryonic stage in pigs.

3. Materials and Methods

Ethics statement

The experimental use of pigs was approved by the Institute of Laboratory Animal Resources, Seoul National University (SNU-140328-2).

BLAST search

The identification of the human *OCT4* and porcine *OCT4A* mRNA homologue region in the porcine genome was performed following a previous report (Hwang *et al.* 2013a). Briefly, the sequence of the entire porcine genome (*Sus scrofa* genome, version 10.2) was compared to the human *OCT4* sequence (GenBank accession No: 5460) using BLAST. The regions with BLAST hit scores of over 100 in the porcine *OCT4* genomic region were considered to be sequence homologue regions. The first exon of the porcine *OCT4* sequence (GenBank accession No: NM_001113060.1) was compared to the whole porcine genome and the hit region on the other chromosome was considered to be the porcine *OCT4A* homologue region. The identified genomic sequences were applied to porcine *OCT4A*-specific primer design.

Chemicals

Unless otherwise description, all chemicals were get from Sigma-Aldrich Corp. (St. Louis, MO, USA).

Sample preparation

Adult tissue Six adult tissues, liver, lung, kidney, heart, spleen, and testis, obtained from purchased a crossbred (Landrace x Large White x Duroc) piglet (d=1, wt=1.5 Kg, Jacob farm, Eumseong, Korea) were ground, frozen by liquid nitrogen, and preserved at - 70°C until used for an experiment.

***In vitro* fertilized embryo** *In vitro* fertilized embryos were obtained following the method described in a previous study (Hwang *et al.* 2013b). Briefly, the pre-pubertal gilt ovaries donated by the Sooam Biotech Research Institute (Seoul, Korea) were used in this experiment. The follicular fluid and cumulus-oocyte complexes (COCs) was aspirated with an 18-gauge needle, and pooled to obtain sediments. Sediments were washed with TL-Hepes-PVA medium (Funahashi *et al.* 1997) and the oocytes with compact cumulus cells and granulated cytoplasm were selected for *in vitro* maturation. The washed COCs were cultured in tissue culture medium (TCM-199; Life Technologies, Rockville, MD) containing 10 ng/ml epidermal growth factor (EGF), 1 µg/ml insulin, and 10% PFF for 44 h at 39°C in 5% CO₂ and 100% humidity conditions. The COCs were treated with hormones, 4

IU/ml eCG and hCG (Intervet, Cambridge, UK), only for the first 22 h and then the COCs were matured in hormone-free conditions. After 44 h of maturation, cumulus cells were detached from the oocytes using 0.1% (w/v) hyaluronidase with gentle pipetting. After removing the cumulus cell, about 20 mature oocytes were placed into 40- μ l micro drops of modified tris-buffer medium (mTBM) (Abeydeera and Day 1997), and covered with 39°C mineral oil. Fresh semen in Duroc (provided by the DARBI A.I. Center in Jochiwon, Korea) was washed by centrifuge with 0.1% (w/v) bovine serum albumin (BSA) -supplemented Dulbecco's phosphate buffered saline (DPBS) and introduced by micro-drops for a final concentration of 1×10^5 cells/ml. The semen and oocytes were co-incubated for 6 h and the attached sperm cells were removed from the oocyte by repetitive pipetting. Sperm-removed fertilized embryos were moved to the culture medium, porcine zygote medium 3 (PZM-3) (Yoshioka *et al.* 2002). The cleaved and non-cleaved embryos were divided after 48 h of culture. Embryos of each stage were pooled after zona pellucida were removed by acid Tyrode's solution.

RNA extraction

Oocyte and preimplantation embryo Pooled oocytes (n=40), each stage of pooled *in vitro* fertilized embryos (2-cell, n=20; 4-cell, n=20; 8-cell, n=20; morula, n=10; and blastocyst, n=5), and a single blastocyst were assessed using the Dynabeads® mRNA DIRECT™ Kit (Invitrogen, Carlsbad, CA, USA) following

the manufacturer's instructions. The zona pellucida was removed by acidified Tyrode's solution before mRNA extraction.

Somatic tissue RNA extraction from adult tissue samples was conducted with TRIzol® Reagent (Invitrogen, Carlsbad, CA) following the manufacturer's instructions. The RNA from tissues was treated with TurboTM DNase I (Applied Biosystems, Foster City, CA, USA) following the manufacturer's instructions.

Reverse transcription

DNase-treated RNA from adult tissue (1 µg), pooled oocytes and embryos, and a single blastocyst was reverse-transcribed using a High Capacity RNA-to-cDNATM Kit (Applied Biosystems, Foster City, CA, USA) in 20 µl of final volume according to the manufacturer's instructions. Reverse-transcriptase non-treated RNA samples were used as negative controls for genomic DNA contamination.

Genomic DNA extraction

Genomic DNA from ground liver (5 mg) was extracted using a G-spinTM Genomic DNA Extraction Kit (iNtRON Bio Technology, Seongnam, Korea) following the manufacturer's instructions. Obtained gDNA was assessed using the porcine *OCT4A*-specific primer test.

PCR

Extracted gDNA and reverse-transcribed cDNA was subjected to PCR with the listed primers (Table V.1). The amplification was carried out using a 2x PCR master mix solution (i-MAX II) (iNtRON Bio Technology, Seongnam, Korea) containing 1 µl of cDNA or 100 ng of gDNA samples and 0.5 µM concentrations of each primer set in 10-µl final reaction volumes with the conditions described in Table V.2. Amplicons were loaded on the 1.5% agarose gels stained with ethidium-bromide (EtBr).

5'- and 3'- RACE PCR

To identify the transcription starting site (TSS) and last sequence of porcine *OCT4* variant transcripts, rapid amplification of cDNA ends (RACE) PCR was carried out using the 5'- and 3'- RACE Core Set (Takara Bio, Otsu, Japan) according to manufacturer's instructions. RNA from pooled *in vitro* fertilized blastocysts (n=20) was assessed by both types of RACE PCR, as follows.

5'-RACE PCR RNA was reverse transcribed with the primer, 5'-AAGGCAGAGGACATG-3', which phosphorylates the first sequence for ligation. Prepared circular forms of cDNA sample were subjected to PCR with 1 µM of each primer, forward: 5'- GGAAAGGGGGTGGGG-3' and reverse: 5'-TCTGTGGGGGTATCTGTAGGC-3'. Nested PCR was performed with 1 µl of

PCR product and 1 μ M of each primer, forward: 5'-GTGGGGTGGGTAAGTGGTA-3' and reverse: 5'-TTATTTGAAGGTCATTACTCCAGA-3'. Detailed PCR conditions are as follow. The first round of PCR was performed according to the following conditions: 1 cycle of 95°C for 7 min; 55°C for 5 min; 72°C for 40 min; 25 cycles of 95°C for 30 sec, 55°C for 1 min 30 sec, and 72°C for 1 min 30 sec; and 1 cycle of 95°C for 40 sec, 55°C for 1 min 30 sec, and 72°C for 15 min. The second round of PCR was carried out as follows: 1 cycle of 95°C for 7 min; 40 cycles of 95°C for 30 sec, 55°C for 1 min, and 72°C for 1 min 30 sec; and 1 cycle of 72°C for 10 min.

3'-RACE PCR RNA was reverse transcribed using adopter oligomer supplied by the kit. PCR and nested PCR were carried out using the forward primer pair, 5'-AGAACCGAGTGAGAGGCAAC-3' and 5'-GCCCAAAGCCCACTCTG-3', respectively. The suggested reverse primer from the 3'-RACE Core Set (Takara Bio, Otsu, Japan) was used for both amplifications. Detailed PCR conditions are described as follow. The first round of PCR was performed as follows: 1 cycle of 95°C for 5 min; 55°C for 5 min; 72°C for 30 min; and 25 cycles of 95°C for 30 sec, 55°C for 1 min, and 72°C for 1 min 30 sec; and 1 cycle of 95°C for 40 sec, 55°C for 1 min 30 sec, and 72°C for 15 min. Nested PCR was performed according to the following conditions: 1 cycle of 95°C for 7 min followed by 40 cycles of 95°C for 30 sec, 55°C for 1 min, and 72°C for 1 min; and 1 cycle of 72°C for 15 min.

Table V.1. Primer sequence used for RT-PCR			
Target gene	Primer	Sequence (5'→3')	Location
<i>OCT4</i> *	AF1**	<u>CCGAGCCCTGTGCCGCC</u>	Exon 1A
	AF2**	<u>GCTGGAGCCGAACCCCGAGG</u>	Exon 1A
	AR1**	<u>CCCCAAA</u> GTGAGCCCCACATCG	Exon 2
	AR2**	CACCTTCCCAAAGAGAA <u>CCCCCAA</u>	Exon 2-3 [‡]
	BF1	GAGCGGCCCTAGAAAAGGCA	Intron 1
	BF2	ATTGCTGTGGCTGTGTTTTG	Intron 1
	BF3	AGGCTGGGGCCTCTTTCCAC	Exon 1B
	R1	CCCCAAAGTGAGCCCCACA	Exon 2
	R2	TCGTTGTTGTCTAGCTTCCTCCAC	Exon 3
	R3	CCCCATAGCCTGGGGTACCAA	Exon 5
<i>OCT4A</i>	AF2	GCTGGAGCCGAACCCCGAGG	Exon 1A
	AR2	CACCTTCCCAAAGAGAAACCCCAAA	Exon 2-3 [‡]
<i>OCT4B</i>	OCT4B_F	GCCTTTTTAAATGCAGTCCCAGG	Exon 1B
	[†] OCT4B_R	TGAACACCTTCCCAAAGAGAAC	Exon 2-3 [‡]
<i>OCT4B1</i>	[†] OCT4B1_F	CCTCACCTTGCTTCTTTCCA	Exon 2B
	OCT4B1_R	CTCTGCCTTGTCATATCTCCTG	Exon 3-4 [‡]
<i>ACTB</i>	ACTB_F	GTGGACATCAGGAAGGACCTCTA	Exon4
	ACTB_R	ATGATCTTGATCTTCATGGTGCT	Exon 4-5 [‡]

*Primer pairs were used for *OCT4A*-specific primer design (AF1, AF2, AR1, and AR2) and *B* and *B1* forms of *OCT4* variant identification

**Bold and underlined sequences on the primer sequences indicate the nucleotides that were different from the *OCT4A* pseudogene (Genbank ID : XR_130606.1)

[†]Indicates primers specifically targeting *B1* and *B*.

[‡]Primers designed to overhang between two marked exons.

Table V.2. Amplification condition for RT-PCR

Target transcript	Ampification condition	Hold 1 (7 min)	Hold 2 (40 Cycles)						Hold 3 (10 min)	Product size (bp)
			Denature		Annealing		Extension			
	Primer*		Temp.	Duration	Temp.	Duration	Temp.	Duration		cDNA
OCT4A	AF1-AR1	95°C	95°C	20 sec	Grad**	20 sec	72°C	30 sec	72°C	182 bp
	AF1-AR2	95°C	95°C	20 sec	Grad**	20 sec	72°C	30 sec	72°C	200 bp
	AF2-AR1	95°C	95°C	20 sec	Grad**	20 sec	72°C	30 sec	72°C	133 bp
	AF2-AR2	95°C	95°C	20 sec	Grad**/68°C	20 sec	72°C	30 sec	72°C	150 bp
OCT4B and OCT4B1	BF1-R2	95°C	95°C	20 sec	63°C	30 sec	72°C	1 min	72°C	N.D
	BF2-R2	95°C	95°C	20 sec	63°C	30 sec	72°C	1 min	72°C	N.D
	BF3-R2	95°C	95°C	20 sec	63°C	30 sec	72°C	1 min	72°C	331bp
	BF3-RE3	95°C	95°C	20 sec	63°C	30 sec	72°C	1min	72°C	459bp / 770 bp
	BF3-BE5	95°C	95°C	20 sec	64°C	30 sec	72°C	1min 30sec	72°C	794 bp / 1,056 bp
OCT4B		95°C	95°C	20 sec	57°C	25 sec	72°C	30 sec	72°C	149 bp
OCT4B1		95°C	95°C	20 sec	60°C	25 sec	72°C	30 sec	72°C	179 bp
ACTB		95°C	95°C	20 sec	63°C	30 sec	72°C	30 sec	72°C	131 bp

*0.5 µM of each primers was used for PCR.

**The annealing temperature was examined by using gradient PCR. The amplification using AF1-AR1, AF1-AR2, and AF2-AR1 pairs were examined at 60°C to 72°C condition with 2°C intervals.

Quantitative RT-PCR

Quantitative PCR was performed with DyNAmo HS SYBR Green qPCR kit (Thermo Scientific, Rockford, IL, USA) and the ABI 7300 Real-Time PCR system (Applied Biosystems, Foster City, CA, USA). For amplification, 0.1 μ M of each primer listed in Table V.3 and 0.5 μ l of cDNA samples from pooled embryos or a single blastocyst were added to the 10- μ l reaction mixtures. The reactions were carried out under the following conditions: 1 cycle of 95°C for 10 min; 40 cycles of 95°C for 15 sec, at annealing temperatures listed in Table V.3 for 60 sec. The dissociation-curve was analyzed to confirm the specificity of the PCR products and the amplified products were gel-loaded. For relative quantity comparison, *ACTB* was used as a control gene.

Absolute quantification of the transcripts

To determine the transcript copy number in each stage of embryo, standard curves of the target genes were developed. The *OCT4* variant amplicons were purified by using the MEGAquick-spin™ Total Fragment DNA Purification Kit (iNtRON Bio Technology, Seongnam, Korea) and the concentration of each purified amplicon was measured (Table V.4). Each purified PCR product was serially diluted 10-fold. Diluted aliquots were assessed by quantitative PCR and Ct values of the 5-step serial dilution (10^{-5} to 10^{-9} dilution) were plotted to show the

log of dilution rate of each PCR product with $R^2 > 0.997$ (Figure V.1 and Table V.4). Because all amplification reactions were performed using 0.5 µl of sample, the copy numbers in 0.5 µl of PCR product were calculated and the number of mRNA copies in 0.5 µl of sample were calculated using a developed plot with the following equation:

Copy number of 0.5 µl sample cDNA

$$= \text{Copy number of 0.5 µl PCR product} \times 10^{\frac{(Ct - Y_{\text{Intercept}})}{\text{Slope}}}$$

After calculating the copy number of a 0.5 µl sample of cDNA, the copy number of each variant in an oocyte/embryo was calculated as follows,

One oocyte or embryo copy number

$$= \frac{\text{Copy number of 0.5 µl sample cDNA} \times 40}{\text{Number of cells or embryos in each pooled sample}}$$

The slope of the developed linear regression plot was in the range of -3.6586 to -3.4727 and was applied to calculate the amplification efficiency (E) with the following equation:

$$\textit{Amplification efficiency (E)} = 10^{\left(\frac{-1}{\textit{slope}}\right)}$$

The amplification efficiency ranged from 1.876 to 1.940 (Table V.4).

Cloning and sequencing

Gel-loaded PCR products were purified by using the MEGAquick-spin™ Total Fragment DNA Purification Kit (iNtRON Bio Technology, Seongnam, Korea) to confirm the sequences. Purified DNA fragments were cloned into a PLUG-prime® TA-cloning vector kit (iNtRON Bio Technology, Seongnam, Korea) and transformed into *E. coli* cells (DH5α, Novagen, Madison, WI). Plasmid was prepared using commercial spin columns (DNA-spin™ Plasmid DNA Purification Kit; iNtRON Bio Technology, Seongnam, Korea). Purified plasmid samples were sequenced with an ABI PRISM 3730 DNA Analyzer (Applied Biosystems, Foster, CA, USA).

Table V.3. Primer sequences used for real time PCR			
Primer	Sequence (5' → 3')*	Size	Tm
<i>OCT4A</i>	F : GCTGGAGCCGAACCCCGAGG R : CACCTTCCCAAAGAGAACCCCCAAA	150 bp	68 °C
<i>OCT4B</i> *	F : GCCTTTTAAATGCAGTCCCAGG R : TGAACACCTTCCCAAAGAGAAC	149 bp	57 °C
<i>OCT4B1</i>	F : CTTTCCACCCACCTCCCCA R : GGTCTCTGCCTTGCATATCTCCT	169 bp	60 °C
<i>SOX2</i>	F : GCCAGAAGAGGAGGGAAGC R : GCGAGGAAAATCAGACGAAGA	91 bp	60 °C
<i>NANOG</i>	F : CTCTCCTCTTCCTTCCTCCA R : TTCCTCCTTGTCTGTGCTCTTC	115 bp	63 °C
<i>CDX2</i>	F : CAGCGGCGGAACCTGTG R : ACTCGGTATTTGTCTTTCGTCCTG	92 bp	63 °C
<i>ACTB</i>	F : GTGGACATCAGGAAGGACCTCTA R : ATGATCTTGATCTTCATGGTGCT	131 bp	63 °C

*The reverse primer of *OCT4B* was designed to overhang between exon 2 and exon 3 to avoid simultaneous detection of the *OCT4B1* transcript together

Table V.4. Porcine <i>OCT4</i> variant amplification plot information								
Target gene	Concentration (PCR product, ng/μl)	Molecular weight (kDa)	Weight / 1 copy (ng) ^a	Copy Number ^b (0.5μl Product)	Amplification plot			
					Slope	Y-intercept	Amplification ^c efficiency	R ²
<i>OCT4A</i>	26.25	92.72588	1.53975E-10	8.5241E+10	-3.5764	-2.4674	1.903753794	0.999
<i>OCT4B</i>	16.25	92.10063	1.53E-10	5.3127E+10	-3.4727	-2.4674	1.940708807	0.9975
<i>OCT4B1</i>	2.45	110.64547	1.83731E-10	6.6674E+09	-3.6586	-1.1193	1.876413702	0.9984

^aOne copy of PCR product (ng) = (Molecular weight of PCR product)/(Avogadro's number) x 10⁹, Avogadro's number = 6.0221413 x 10²³

^bCopy number (0.5μl of PCR product) = 0.5 x (PCR product concentration, ng/1ul) / (one copy of PCR product, ng)

^cAmplification efficiency = 10^(-1/slope)

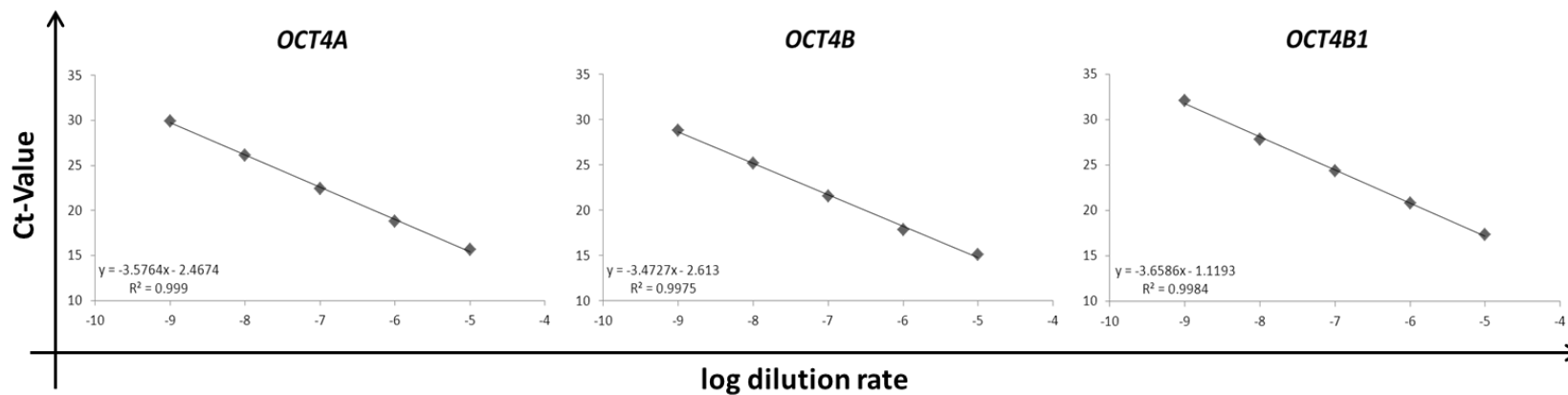


Figure V.1 Amplification plot of each primer pair used for real time PCR. The Ct-value was plotted against the 10-fold serial dilution rate of each PCR product (n=3). The equation of linear plots was applied to calculate the copy number of each sample.

Western blot

The ground pig tissues were lysed with RIPA buffer (Sigma-Aldrich, St. Louis, MO, USA) to extract total protein. Fifty micrograms of extracted protein were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis and transferred to nitrocellulose transfer membrane (Schleicher & Schuell Bioscience, Dassel, Germany). The membrane was blocked with 5% (w/v) skim milk for 1 h at room temperature. Three NTD targeting anti-OCT4 primary antibodies, mouse-anti-OCT4 (1:200, SC-5279, SantaCruz, CA, USA); goat-anti-OCT4 (1:200, SC-8628, SantaCruz, CA, USA); rabbit-anti-OCT4 (1:200, SC-9081, SantaCruz, CA, USA), and CTD targeting goat-anti-OCT4 (1:200, SC-8629, SantaCruz, CA, USA) were used after membrane blocking. Rabbit-anti-ACTIN (1:500, A2066, Sigma-Aldrich, St. Louis, MO, USA) and rabbit-anti-GAPDH (1:500, 5174, Cell Signaling Technology, Boston, MA) were used as internal controls. The membrane was treated with primary antibodies overnight at 4°C. After incubation, the membrane was washed 3 times with TBS-T and incubated with goat-anti-mouse IgG-HRP (1:1000, sc-2005, SantaCruz, CA) for SC-5279, donkey-anti-goat IgG-HRP (1:1000, sc-2020, SantaCruz, CA) for SC-8628 and 8629, and goat-anti-rabbit IgG-HRP (1:1000, SC-2004, SantaCruz, CA) for SC-9081, GAPDH, and ACTIN for 1 h at room temperature. The membrane was then washed three times after secondary antibody treatment and the protein expression was detected with enhanced chemiluminescence (Amersham Bioscience, Little Chalfont Bucks, UK) and visualized by ChemiDoc (ChemiDoc™ XRS+ System, Bio-Rad, Hercules,

CA).

Immunocytochemistry

The each stage of embryos and mature oocyte without zona pellucida were fixed by 4% paraformaldehyde for 30 min at 4°C. Fixed samples were permeabilized by 1% Triton-X 100 for 5 min at room temperature and washed three times with DPBS. The embryos and oocytes were blocked using 1% BSA in DPBS for 1 h at room temperature and primary antibodies targeting NTD or CTD of OCT4 (SC-5279, mouse-anti-OCT4 or SC-8629, goat-anti-OCT4, respectively) were added overnight at 4°C. Each primary antibody was diluted 1:100 in DPBS containing 1% BSA. Embryos were washed three times using DPBS with 0.1% Tween-20 before incubating with fluorescent-conjugated secondary antibodies, anti-mouse-IgG (1:500, A11029, Invitrogen Carlsbad, CA, USA) and anti-goat-IgG (1:200, SC-362284, SantaCruz, CA, USA), which were diluted in blocking solution, for 1 h at room temperature. The samples were washed three-times using 0.1% Tween-20 in DPBS and the nucleus was stained by 0.1% Hoechst 33342 (Molecular Probes, Eugene, OR, USA) for 10 min. After washing three times with DPBS, samples were mounted on the slide glass. Stained samples were visualized using an LSM 700 Laser Scanning Microscope (Carl Zeiss, Oberkochen, Germany) and captured images were processed with ZEN 2012 Light Edition software (Carl Zeiss, Oberkochen, Germany).

Statistical analysis

Statistical analysis was carried out using the Graphpad Prism statistical program (Graphpad Software, San Diego, CA). Porcine *OCT4* variant expression in oocytes and developing preimplantation embryos were examined using one-way analysis of variance (ANOVA) and the Tukey test. Differential gene expressions between the two classified groups of individual blastocyst were compared using the unpaired Student *t*-test. Correlation between *OCT4* variant and pluripotent gene expression in a single blastocyst was analyzed by the Pearson correlation coefficient. All data were presented as mean \pm S.E.M and $P < 0.05$ was considered to be significant.

4. Results

Identification of novel porcine *OCT4* variants, *OCT4B* and *OCT4B1*

To examine the presence of porcine *OCT4* novel variants, the human *OCT4* gene sequence was compared to the porcine genome using BLAST. The results showed that the 3' region of the first intron of porcine *OCT4* has similarity with the first exon of the human *OCT4B* transcript (Figure V.2). Following the alignment, primers were prepared to analyze expression of *OCT4* variants in pig (Figure V.3A and Table V.1). Because *OCT4B* was reported to be expressed in human preimplantation embryo (Cauffman *et al.* 2006), the cDNA samples from *in vitro* fertilized blastocysts (n=15) were subjected to RT-PCR. The region from BF3 to R3 was successfully amplified (Figure V.3B) and the PCR product was validated by sequencing (Figure V.4). This result demonstrated that porcine *OCT4B* variants also share the second and last exon of porcine *OCT4A*, as is seen in human and mouse. The additional amplicon, which is approximately 300-bp larger than the expected PCR product, was consistently detected when the reverse primer amplifying downstream of the second intron (R2 and R3) was subjected to PCR (Figure V.3B). The sequence of the amplicon demonstrated that the PCR product spans the second intron of porcine *OCT4A* (Figure V.4) and that the structure of the variant is similar to that of human *OCT4B1* mRNA. Therefore, in this study, the

variant harboring the second intron in its transcript was named as porcine *OCT4B1*. Based on the identified partial sequence of novel porcine *OCT4* variants, RACE PCR was carried out to confirm the full sequence of each variant. Both variants share TSS, which is located 432 bp upstream from the second exon of porcine *OCT4A* (Figure V.3C). The last sequence of both variants was identical to that of the porcine *OCT4A* transcript (Figure V.3D). Identified new porcine *OCT4* variants, *OCT4B* and *OCT4B1* (Genbank accession No. KJ023671 and KJ023672, respectively), have distinct first exons (E1B) compared to the A form, and *OCT4B1* has a cryptic exon (E2B), which is in the intron region in porcine *OCT4A* and *OCT4B*, as it is in human *OCT4B1* (Figure V.3E). These result showed that porcine *OCT4* transcribes additional mRNAs, *OCT4B* and *OCT4B1*, as well as *OCT4A*.

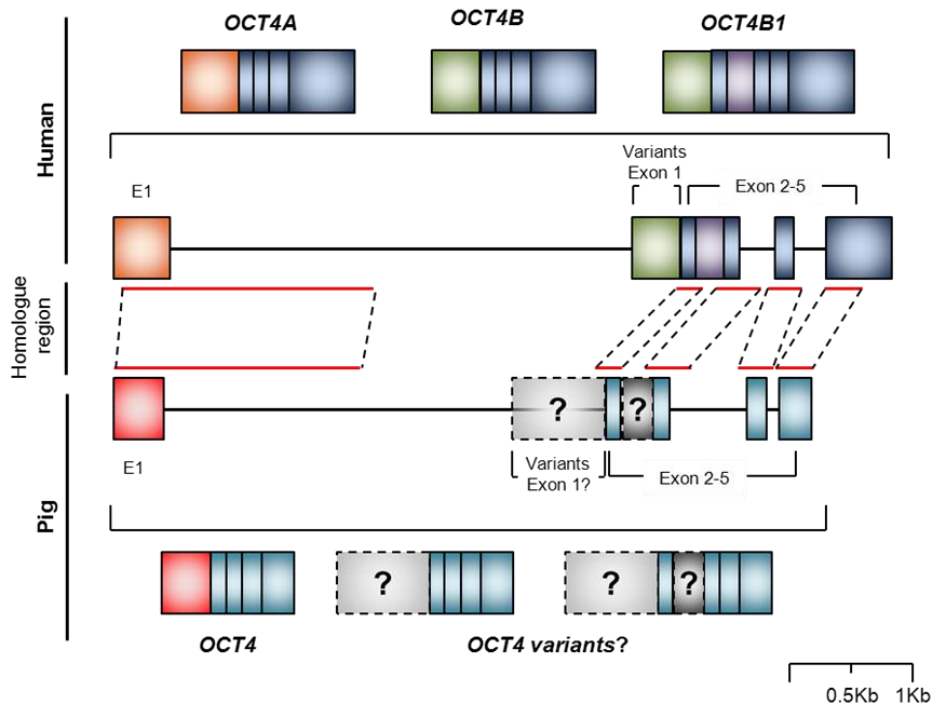


Figure V.2. Identification of candidate coding regions of porcine *OCT4* variants by searching for regions homologous to human *OCT4*. The structure and sequence of homologue sites between human and porcine *OCT4* are represented. Human *OCT4* variants showed differences in the three regions (orange, green, and purple rectangles). The region homologous to the human *OCT4* variants coding region in the porcine genome was searched by BLAST (red lines). Based on the sequence similarity and location of the first exon of human *OCT4B* (green box) and *OCT4B1*-specific exon (purple box), porcine *OCT4B* and *B1*-specific sites were determined (gray boxes with question marks). The diagram is scaled.

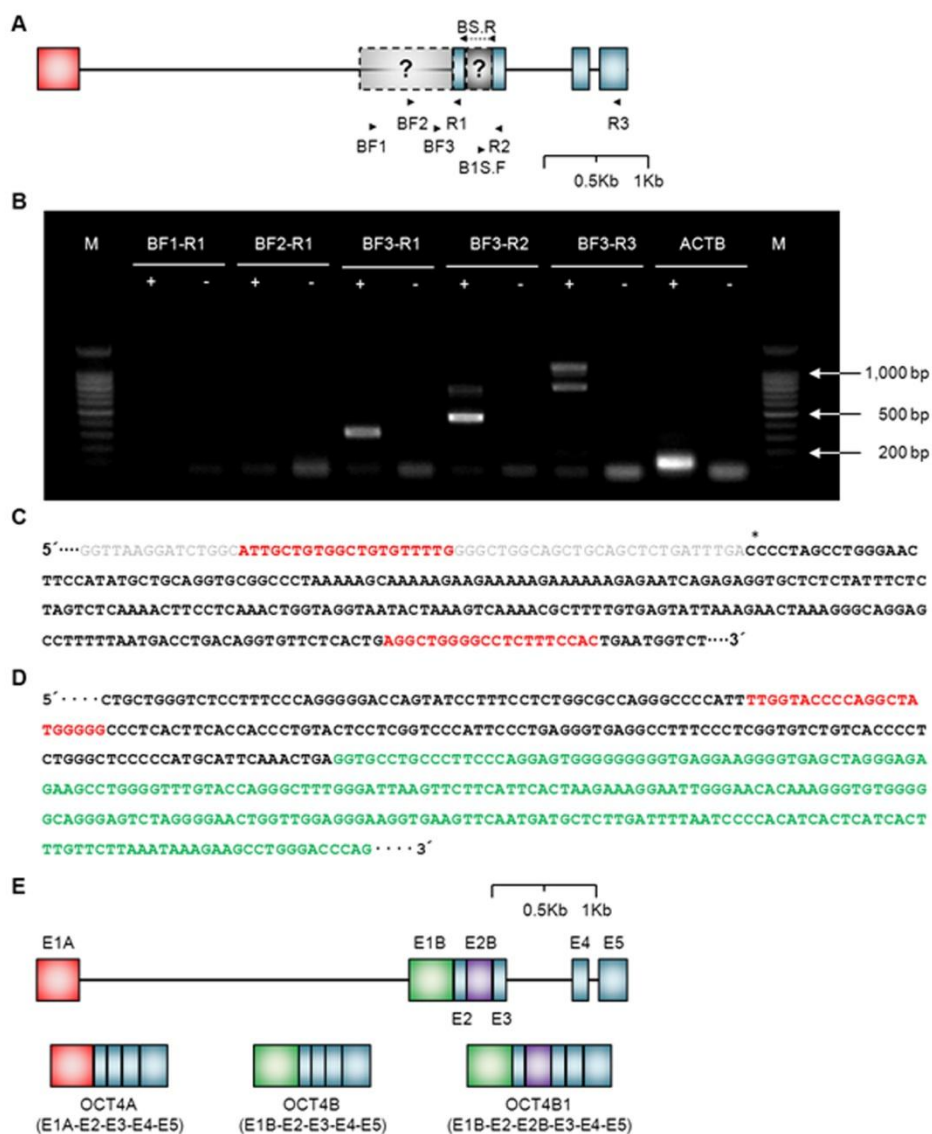


Figure V.3. Identification of porcine *OCT4* variants, *OCT4B* and *OCT4B1*. (A)

Porcine *OCT4* gene structure and anticipated variants. Rectangles indicate exon 1 (red) and exons 2 to 5 (blue). Gray boxes with question marks indicate possible coding regions for *OCT4B* and *OCT4B1* transcripts. Primers are shown and the primer noted as BS.R indicates the primer located between exon 2 and 3 used to distinguish *OCT4B* and *OCT4B1*. The diagram is scaled. (B) Candidate *OCT4* variant expression in porcine blastocyst. The cDNA (+) from pooled blastocyst (n=15) was successfully used to amplify the BF3 to R3 region. Non-reverse transcribed RNA (-) was used for a negative control and M indicates the 100 bp-DNA ladder. (C) Identified transcription starting sites (TSSs). The sequence marked by an asterisk is the first sequence of the novel *OCT4* variant. Red characters indicate BF2 (upper) and BF3 (lower) primers. Black sequences and gray sequences represent expression and non-expression sites, respectively. (D) Identified last sequences. Last sequences were the same as those previously reported for the *OCT4A* mRNA sequence. Red text indicates the R3 primer-complimentary sequences. Black and green characters represent exon 5 and the 3'-UTR, respectively. (E) Identified porcine *OCT4* gene structure and variant mRNA. Boxes and lines indicate exons and introns, respectively. Boxes of red, green, and purple are *OCT4A*, *OCT4B*, and *OCT4B1*-specific regions, respectively. The diagram is scaled.

OCT4	1	TTCTTAGTCTCAAACTTCCTCAAACTGGTAGTAATACTAAAGTCAAAACGCTTTTGTAGTATTAAAGAACTAAAGGGCAGAGCTTTTAAATGAC	100
F3-RE5_Long	1	1
F3-RE5_Short	1	1
Clustal Consensus		1
OCT4	101	CTGACAGGTGTTCTCACTGAGGCTGGGGCTCTTCCACGGAATGGCTCTCTGCAATTGATAGGTGTGTCACCTTCTGGAGTAAATGACCTTCAAAATAAA	200
F3-RE5_Long	82	181
F3-RE5_Short	1	AGGCTGGGGCTCTTCTCACTGGAATGGCTCTCTGCAATTGATAGGTGTGTCACCTTCTGGAGTAAATGACCTTCAAAATAAA	81
Clustal Consensus		181
OCT4B/B1-specific region			
OCT4	201	GTTATTCTGGTGTGTTAGGTGGGTAGCTACAGATACCCACAGACCATTTACTGCTTTGGGGCTAAAGTGGAAAGGGGTGGGTGGGTAAAGTGGTAC	300
F3-RE5_Long	82	181
F3-RE5_Short	82	GTTATTCTGGTGTGTTAGGTGGGTAGCTACAGATACCCACAGACCATTTACTGCTTTGGGGCTAAAGTGGAAAGGGGTGGGTGGGTAAAGTGGTAC	181
Clustal Consensus	82	181
Exon2			
OCT4	301	GTATATGTTCTACATGTCCTCTGCTTTTAAATGCAATCCAGGACATCAAAAGCTTTCAGAAAGATCTGAAACATTTGCCAAGCTCTCAAAACGAG	400
F3-RE5_Long	182	281
F3-RE5_Short	182	GTATATGTTCTACATGTCCTCTGCTTTTAAATGCAATCCAGGACATCAAAAGCTTTCAGAAAGATCTGAAACATTTGCCAAGCTCTCAAAACGAG	281
Clustal Consensus	182	281
OCT4	401	AGAGGATCACCTGGGATATACCCAGGCCGATGTGGGGCTCACTTTGGGGGTCTCTTTGTGAGTCTTCCTCAATATGTTCTGACCCACAGAACACAGG	500
F3-RE5_Long	282	341
F3-RE5_Short	282	AGAGGATCACCTGGGATATACCCAGGCCGATGTGGGGCTCACTTTGGGGGTCTCTTTGTGAGTCTTCCTCAATATGTTCTGACCCACAGAACACAGG	341
Clustal Consensus	282	341
OCT4	501	CCACACAGTGTCCCTAGAAAGAACATCCCAAAATGTGGGCTTTCTCCGCCCAAGGAGGTGAGGGAGGGGGTTCACCCCTTTTATGGGTGAGGGG	600
F3-RE5_Long	382	481
F3-RE5_Short	341	CCACACAGTGTCCCTAGAAAGAACATCCCAAAATGTGGGCTTTCTCCGCCCAAGGAGGTGAGGGAGGGGGTTCACCCCTTTTATGGGTGAGGGG	341
Clustal Consensus		481
Intron 2			
OCT4	601	AGGGGGGAGAAAGATACCCCTACTTGTCTTGGTCTTGGCTTGCACTTCTTGGAGGAGTTCCTGCAATGGGTCCCAAGGAGGGGGAGAGAGGTGGCA	700
F3-RE5_Long	482	581
F3-RE5_Short	341	AGGGGGGAGAAAGATACCCCTACTTGTCTTGGTCTTGGCTTGCACTTCTTGGAGGAGTTCCTGCAATGGGTCCCAAGGAGGGGGAGAGAGGTGGCA	341
Clustal Consensus		581
OCT4	701	GGCTTTAGTCTTTGGGTGCGGAGGGGCTGCTGCTGCTCAGCTTGCTTTCTTCCACCCACTCCCAAGGAAAGTGTTCAGCCAAACGACCATCTGG	800
F3-RE5_Long	582	681
F3-RE5_Short	341	GGCTTTAGTCTTTGGGTGCGGAGGGGCTGCTGCTGCTCAGCTTGCTTTCTTCCACCCACTCCCAAGGAAAGTGTTCAGCCAAACGACCATCTGG	370
Clustal Consensus		681
Exon3			
OCT4	801	CGTTTGAAGCTTTGCACTCACTTTCAAGAACATGTATAAGCTGCGGCCCTGCTGCAAGATGGGTGGAGGAACTGACACAAACGAGAACTCTGCAAG	900
F3-RE5_Long	682	781
F3-RE5_Short	371	CGTTTGAAGCTTTGCACTCACTTTCAAGAACATGTATAAGCTGCGGCCCTGCTGCAAGATGGGTGGAGGAACTGACACAAACGAGAACTCTGCAAG	470
Clustal Consensus	371	470
OCT4	901	AGGTGAGGTGGAGGGGTGCACAGGGGCCGCCCATCTCAGCCCACTCTCAGCCCAATTGCACTGCTTTTGTCCCTGGCTGGCAGCTCTCTTTTGG	1000
F3-RE5_Long	782	783
F3-RE5_Short	471	AG	472
Clustal Consensus	471	..	472
OCT4	1001	AGGCACTGGCTCTCGGTGGGTGGGTGTAAATCTCACTGTTGTTGGAGAAAGTCCAGCTGGGGCACTGTGCTAAAGCTACCTTTCAGATCTTTTC	1100
F3-RE5_Long	783	783
F3-RE5_Short	472	472
Clustal Consensus		472

(Continued)

	1101	1120	1140	1160	1180	1200	1220	1240	1260	1280	1300
OCT4	1101	TTTCCATTATTATTATTATTTTGTCTTTCTTAGGGCCACACCCGACCAATGGAGTTCCAGGCTAGGGTCCAACTGGAGCTATAGTGCACG	1200								
F3-RE5_Long	783	-----	-----	-----	-----	-----	-----	-----	-----	-----	783
F3-RE5_Short	472	-----	-----	-----	-----	-----	-----	-----	-----	-----	472
Clustal Consensus	472	-----	-----	-----	-----	-----	-----	-----	-----	-----	472
	1210	1230	1250	1270	1290	1310	1330	1350	1370	1390	1400
OCT4	1201	CTACGCCAGAGCCACACAGCCATGTGGATCTGAGCCAGCTCTGTGACCTAGCCACAGCTCATGGCAACGGATCTTAACCCACTGAGCAAGC	1300								
F3-RE5_Long	783	-----	-----	-----	-----	-----	-----	-----	-----	-----	783
F3-RE5_Short	472	-----	-----	-----	-----	-----	-----	-----	-----	-----	472
Clustal Consensus	472	-----	-----	-----	-----	-----	-----	-----	-----	-----	472
	1310	1330	1350	1370	1390	1410	1430	1450	1470	1490	1400
OCT4	1301	CAGGGATCAAACTGCAACCTCATGGTTTTAGTCAGATTGTTAATCACTGAGTCACGACAGGACCTCCAGGATCTTCTCTTTTGTAACTGGCCCT	1400								
F3-RE5_Long	783	-----	-----	-----	-----	-----	-----	-----	-----	-----	783
F3-RE5_Short	472	-----	-----	-----	-----	-----	-----	-----	-----	-----	472
Clustal Consensus	472	-----	-----	-----	-----	-----	-----	-----	-----	-----	472
	1410	1430	1450	1470	1490	1510	1530	1550	1570	1590	1600
OCT4	1401	GGTGACACATCCAGCCAGTGTCCGAGCAGTTTGCTCTTAACTGACTGACCTGGGTTGCCCTGCTCAGCAGTTTGAGGCTTGCCCTTCCCATTTCC	1500								
F3-RE5_Long	783	-----	-----	-----	-----	-----	-----	-----	-----	-----	783
F3-RE5_Short	472	-----	-----	-----	-----	-----	-----	-----	-----	-----	472
Clustal Consensus	472	-----	-----	-----	-----	-----	-----	-----	-----	-----	472
Exon4											
	1510	1530	1550	1570	1590	1610	1630	1650	1670	1690	1700
OCT4	1501	TGCTTAAATTTCTCTCCCATCTGCCAGATATGCAAGCAGAGACCTGTGCAAGCCCGAAGAGAAAGCGACAAGTATGAGAACCGAGTGAGAG	1600								
F3-RE5_Long	783	-----	-----	-----	-----	-----	-----	-----	-----	-----	853
F3-RE5_Short	472	-----	-----	-----	-----	-----	-----	-----	-----	-----	542
Clustal Consensus	472	-----	-----	-----	-----	-----	-----	-----	-----	-----	542
	1610	1630	1650	1670	1690	1710	1730	1750	1770	1790	1800
OCT4	1601	GCAACCTGGAGAGCATGTTCTGCACTGCCCCAAAGCCCACTCTGCACAGATCAGCCAGATGCCCCAGCAGCTGGGCTAGAGAGGATGT	1700								
F3-RE5_Long	854	-----	-----	-----	-----	-----	-----	-----	-----	-----	944
F3-RE5_Short	543	-----	-----	-----	-----	-----	-----	-----	-----	-----	633
Clustal Consensus	543	-----	-----	-----	-----	-----	-----	-----	-----	-----	633
	1710	1730	1750	1770	1790	1810	1830	1850	1870	1890	1900
OCT4	1701	TGCCCCATCCCCGTGGGCTCCCACTCTCCCCACCCCTCCAGAGCTTATGATCCCAATGCTCTCTCCCTGAACTCTGCTCTCAAGT	1800								
F3-RE5_Long	944	-----	-----	-----	-----	-----	-----	-----	-----	-----	1054
F3-RE5_Short	633	-----	-----	-----	-----	-----	-----	-----	-----	-----	743
Clustal Consensus	633	-----	-----	-----	-----	-----	-----	-----	-----	-----	743
	1810	1830	1850	1870	1890	1910	1930	1950	1970	1990	2000
OCT4	1801	TGGTTCTGCAACCTGCGCCAGAGGGCAACGATCAAGCACTGACTATTCCCAACGAGAGGATTTGAGGCTGCTGGGTCTCTTTCCCAAGGGGACCA	1900								
F3-RE5_Long	955	-----	-----	-----	-----	-----	-----	-----	-----	-----	1054
F3-RE5_Short	644	-----	-----	-----	-----	-----	-----	-----	-----	-----	743
Clustal Consensus	644	-----	-----	-----	-----	-----	-----	-----	-----	-----	743
Exon5											
	1910	1930	1950	1970	1990	2010	2030	2050	2070	2090	2100
OCT4	1901	TATCTTTCTCTTGGCGCCAGGGCCCATCTGGGTACCCAGGCTATGGGGCCCTCACTTCACCCACCTGTAATCTCTGGTCCCATTCCTCGAGGGTGA	2000								
F3-RE5_Long	1055	-----	-----	-----	-----	-----	-----	-----	-----	-----	1105
F3-RE5_Short	744	-----	-----	-----	-----	-----	-----	-----	-----	-----	794
Clustal Consensus	744	-----	-----	-----	-----	-----	-----	-----	-----	-----	794
	2010	2030	2050	2070	2090	2100					
OCT4	2001	GGCTTTTCTCTGGTGTCTGT	2021								
F3-RE5_Long	1105	-----	-----	-----	-----	-----					
F3-RE5_Short	794	-----	-----	-----	-----	-----					
Clustal Consensus	794	-----	-----	-----	-----	-----					

Figure V.4. Long and short BF3-RE5 PCR product sequence alignments to the porcine *OCT4* gene. Long and short fragments of the BF3-RE5 amplicon sequence shown contain the expected *OCT4B* and *OCT4B1*-specific exon sequences as well as exon 2 to exon 5. The larger PCR product contains the same second intron as the human *OCT4B1* transcript. Arrows indicate BF3 and RE5 primers.

Designing an *OCT4A*-specific primer pair

Prior to examine *OCT4* variants expression in various tissues and embryonic stage, *OCT4A* specific primer pairs were designed first because the *OCT4A*-pseudogenes would complicate attempts to measure accurate amounts of *OCT4A* mRNA (Liedtke *et al.* 2007). As human *OCT4* pseudogenes don't have introns, and two genomic sequences similar to the porcine *OCT4A* specific exon, E1A, and E2, were searched in here (Figure V.5), gDNA PCR was performed to know whether the designed primer pairs don't amplify porcine *OCT4A* pseudogenes or not. Production of amplicon with expected size resulted from gDNA PCR means the primer pair detect pseudogene, and concluded that the pair is not suitable for distinguishing porcine *OCT4A* and its pseudogenes. Examined primer sets didn't target pseudogene except AF1-AR1 pair, and the AF2-AR2 pair was determined to be the most desirable primer pair because it detected the fewest number of non-specific amplicons among them (Figure V.6A). The AF2-AR2 primer pair was used to examine *OCT4A* expression in various porcine tissues and successful amplification was detected only in testis (Figure V.7A) without non-specific products detected from the gDNA amplification. The amplicon was sequenced and no *OCT4* pseudogenes were detected (Figure V.6B). Therefore, the designed primer pair in this study is suitable for specifically detecting porcine *OCT4A* mRNA that allows the accurate quantification of *OCT4A* mRNA with a reproducible and reliable discrimination among porcine *OCT4* variants.

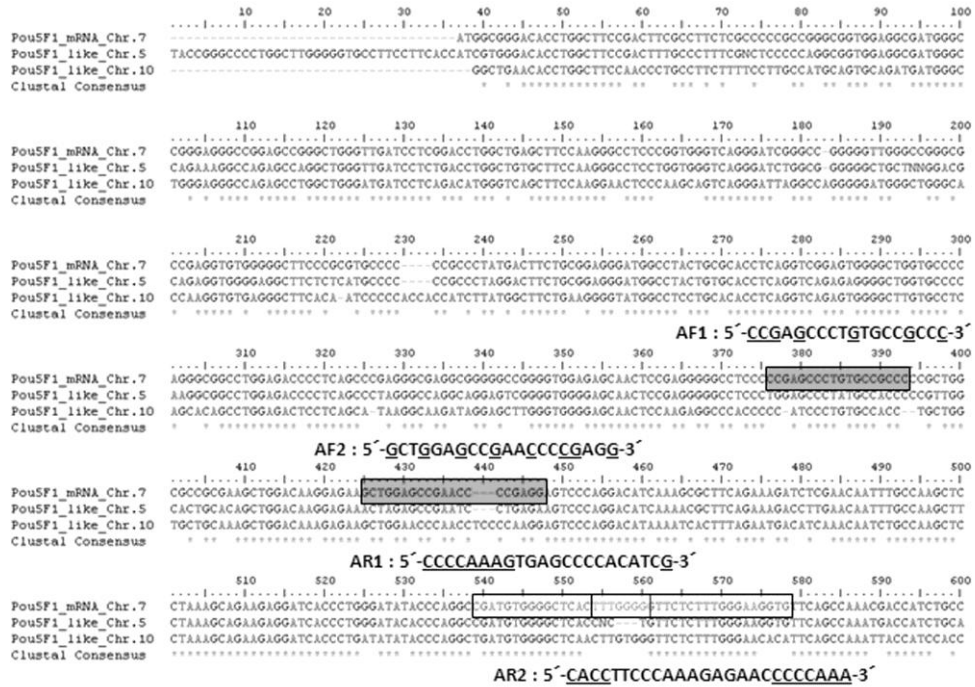


Figure V.5. Porcine *OCT4A* mRNA sequence alignment to *OCT4A* homologue sequences. To design *OCT4A*-specific primers, the annotated *OCT4A* mRNA (Genbank accession No: NM_001113060.1) sequence and its homologue sequences (Genbank accession Nos: XR_130606.1 and XM_003480877.1, respectively) were aligned using CLSUTALW. With the results, forward primers (gray boxes) and reverse primers (white boxes) were designed. Underlined characters indicate the parts of the sequences that differed with the aligned porcine *OCT4A* homologue sequences.

Figure V.6. Specificity validation of porcine *OCT4A*-specific primer by gDNA amplification and sequencing. (A) Genomic DNA amplification with designed *OCT4A* primer. Each number on the lane indicates annealing temperature during the amplification reaction. M indicates 100 bp DNA ladder. (B) AF2-AR2 amplicon sequence alignment to *OCT4A*. The amplicon from embryos was cloned and sequenced. Fifteen cloned colonies were sequenced and all clones showed porcine *OCT4A* sequence, not its pseudogene. Arrows indicate forward (AF2) and reverse (AR2) primers.

Expression of *OCT4* variants in adult tissue

The expression of porcine *OCT4* variants was examined in six male somatic tissues - liver, lung, kidney, heart, spleen, and testis. Porcine *OCT4B* was expressed in all tissues but *OCT4B1* was detected only in testis (Figure V.7A). The PCR products were sequenced and the amplicon was confirmed as the targeted product. After examining mRNA expression of each variant in tissues, the open reading frames (ORFs) from *OCT4B* and *OCT4B1* transcript were analyzed (Table V.5) using the RevTrans ORF finder (Wernersson and Pedersen 2003). Three and eight ORFs starting with the ATG and CTG start codons, respectively, were identified from the *OCT4B* template sequence (Table V.5). The ORF from *OCT4B1* produced a truncated protein because of the stop codon in the E2B, as is seen in human *OCT4B1*. As an antibody targeting OCT4B1 was not available commercially, only OCT4B expression was examined by western blot in this study. The protein detected by OCT4 C-terminal targeting antibody (SC-8629) was slightly less than 25 Kda (Figure V.7B). Compared to the identified ORF information, the detected protein would be OCT4B-201AA (22.38 Kda). However, OCT4A was not detected in testis tissues even though various OCT4A-specific antibodies as well as CTD targeting OCT4 antibody were used (data not shown). This absence could be a result of a heterogenic cell population in assessed testis tissue and a low proportion of OCT4A-positive cells. The results demonstrate that porcine *OCT4B* has protein-coding variants, as is seen in other species, and it is broadly expressed in adult tissues, contrary to *OCT4A* and *OCT4B1* of which mRNA were expressed only in

testis.

Comparative analysis of *OCT4* variants mRNA in porcine oocytes and preimplantation embryos

After identifying the novel porcine *OCT4* variants, *OCT4B* and *B1*, the expression quantity of each variant in mature oocytes and developing preimplantation embryos was examined (Figure V.8). All variants were expressed in oocytes and developing embryos. The amounts of *OCT4A* and *OCT4B* transcripts per one embryo and oocyte were significantly higher in the blastocyst stage compared to other embryonic stages (Figure V.8A). After analyzing the expression of each variant in oocytes and developing embryos, the proportions of variants in each embryonic stage was examined (Figure V.8B). The *OCT4B* and *OCT4B1* mRNA comprised approximately 7% to 19% and less than 1% to 7%, respectively, among porcine *OCT4* variants in developing embryo. The proportion of *OCT4A* transcript was significantly higher compared to proportions of other variants in all examined preimplantation embryonic stages. This finding indicates *OCT4A* transcript is dominantly expressed in porcine embryonic stage rather than newly identified *OCT4B* and *OCT4B1* mRNA.

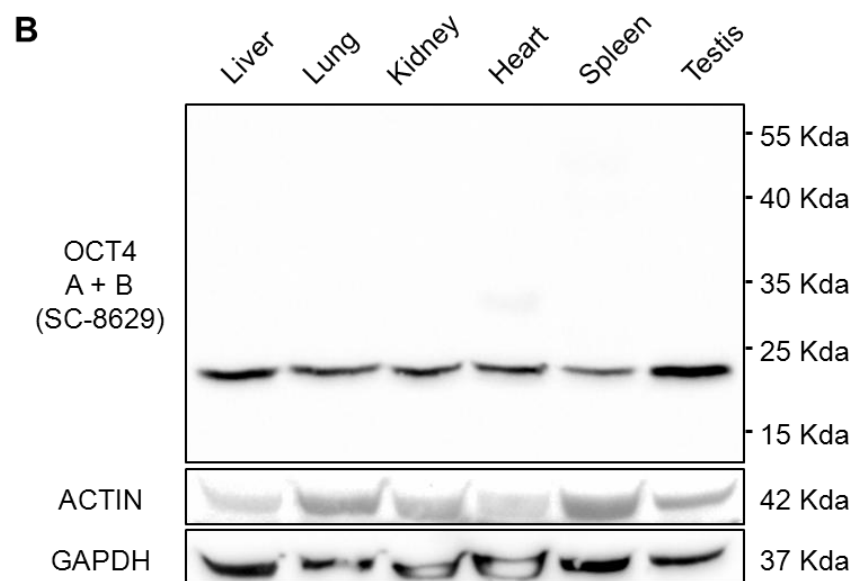
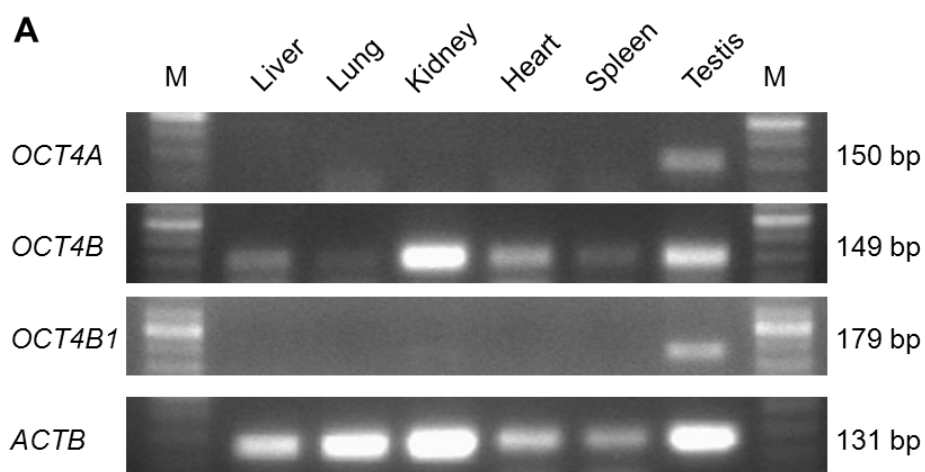


Figure V.7. Expression of porcine *OCT4* variants in somatic tissue. (A)

Discriminative analysis of porcine *OCT4* variant transcript expression in adult tissue. Expression of each variant in adult tissues was examined discriminatively. M indicates the 100 bp DNA ladder. Each PCR product was confirmed by sequencing. (B) OCT4B expression in somatic tissue. Fifty micrograms of whole protein from each tissue were assessed by western blot. The antibody that detects the C-terminal OCT4 domain was used to examine the B form of OCT4 production in porcine tissues. ACTIN and GAPDH were used as internal controls.

Table V.5. Candidate OCT4B protein isoform predicted by transcript translation ^a					
Template	Start codon	Isoform ^b	Starting nt No. ^c	Stop codon location	Molecular weight (Kda)
<i>OCT4B</i>	ATG	OCT4B-227AA	427	Exon5	25.45
		OCT4B-164AA	616		18.31
		OCT4B-113AA	769		12.29
	CTG	OCT4B-201AA	505		22.38
		OCT4B-161AA	625		17.97
		OCT4B-159AA	634		17.6
		OCT4B-158AA	637		17.49
		OCT4B-145AA	676		15.92
		OCT4B-117AA	760		12.64
		OCT4B-112AA	775		12.03
		OCT4B-105AA	796		11.26
		OCT4B-38AA	967		4.85
		OCT4B-29AA	1,024		2.95
<i>OCT4B1</i>	ATG	OCT4B1-77AA	427	Exon2B	8.14

^aThe porcine OCT4B isoform was predicted using the ORF finder tool, RevTrans 2.0 (Rasmus and Anders, 2003)

^bEach isoform was termed as follows; OCT4B/OCT4B1-"The number of amino acid"

^cTranslation starting site was represented by the nucleotide (nt) order of *OCT4B* and *OCT4B1* mRNA sequence (GenBank ID : KJ023671 and KJ023672)

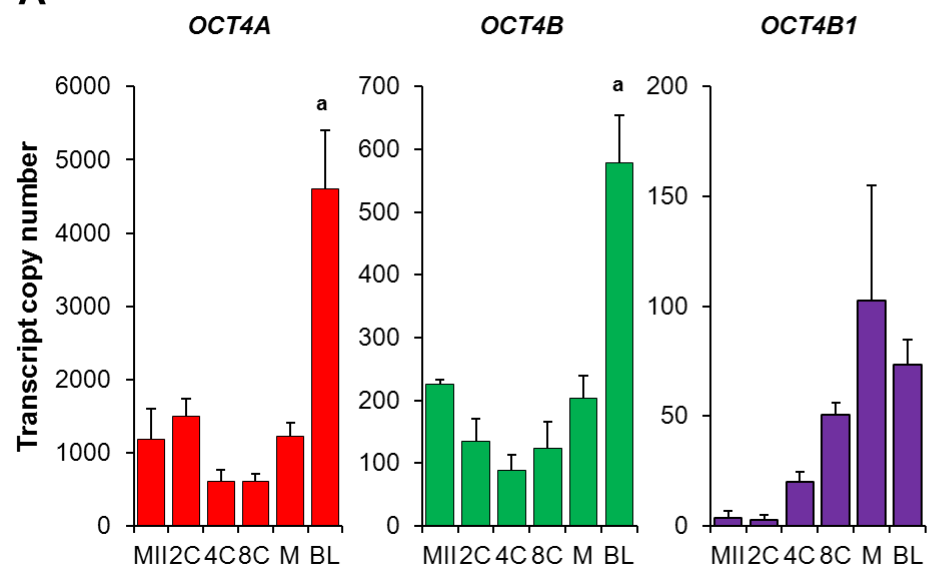
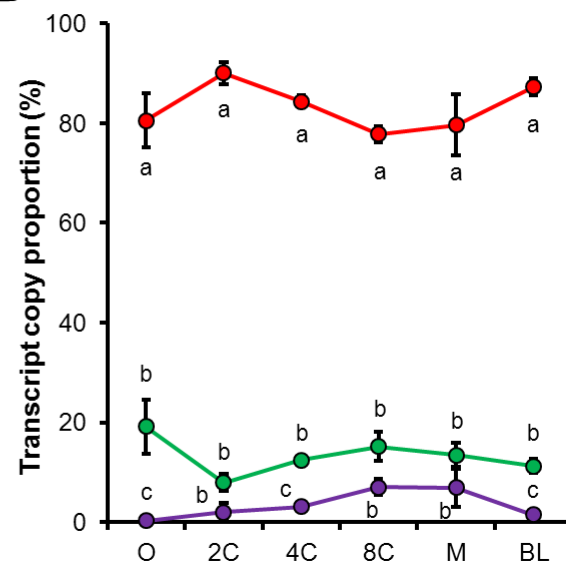
A**B**

Figure V.8. Quantitative comparison of mRNA of each porcine *OCT4* variant in mature oocytes and developing preimplantation embryos. (A) The mRNA copy number of each variant in developing embryo. The absolute amount of each variant was examined discriminatively in mature oocyte (MII, n=40) and 2-cell (2C, n=20), 4-cell (4C, n=20), 8-cell (8C, n=20), morula (M, n=10), and blastocyst (BL, n=5) preimplantation embryonic stages. Copy number of each variant in an oocyte and embryo was calculated by an equation described in the materials and methods section. The data are presented as mean \pm S.E.M. Different characters indicate significant differences ($P<0.05$). (B) Each variant proportion in oocytes and developing embryos. The proportion of each variant was analyzed by comparing the copy numbers of each variant in oocytes (O) and preimplantation embryos (2-cell, 2C; 4-cell, 4C; 8-cell, 8C; morula, M; blastocyst, BL). The data are presented as mean \pm S.E.M. Different characters indicate significant differences ($P<0.05$).

Porcine OCT4A and OCT4B expression and location in oocytes and preimplantation embryos

The expression and locations of OCT4A and OCT4B were examined in mature oocytes and developing embryos (Figure V.9). Because of a lack of antibody to OCT4B1, it was excluded from these experiments. OCT4A was detected in blastocyst (Figure V.9A) but OCT4B was detected in all embryonic stages and in mature oocytes (Figure V.9B). Interestingly, each protein had distinct locations in embryo. OCT4A was detected only in nucleus while OCT4B was found expressed in cytoplasm and membrane. However, the antibody targeting OCT4 CTD, which was expected to detect both the A and B forms of OCT4, showed unclear specificity for detection of OCT4A because the certain expression in nucleus was not observed in blastocyst (dashed area in Figure V.9B), which showed positive expression of OCT4A (Figure V.9A). These results indicate that the OCT4B protein is expressed in differential stages of and locations in embryo compared to OCT4A expression.

Porcine OCT4A and OCT4B expression and location in oocytes and preimplantation embryos

The expression and locations of OCT4A and OCT4B were examined in mature oocytes and developing embryos (Figure V.9). Because of a lack of antibody to

OCT4B1, it was excluded from these experiments. OCT4A was detected in blastocyst (Figure V.9A) but OCT4B was detected in all embryonic stages and in mature oocytes (Figure V.9B). Interestingly, each protein had distinct locations in embryo. OCT4A was detected only in nucleus while OCT4B was found expressed in cytoplasm and membrane. However, the antibody targeting OCT4 CTD, which was expected to detect both the A and B forms of OCT4, showed unclear specificity for detection of OCT4A because the certain expression in nucleus was not observed in blastocyst (dashed area in Figure V.9B), which showed positive expression of OCT4A (Figure V.9A). These results indicate that the OCT4B protein is expressed in differential stages of and locations in embryo compared to OCT4A expression.

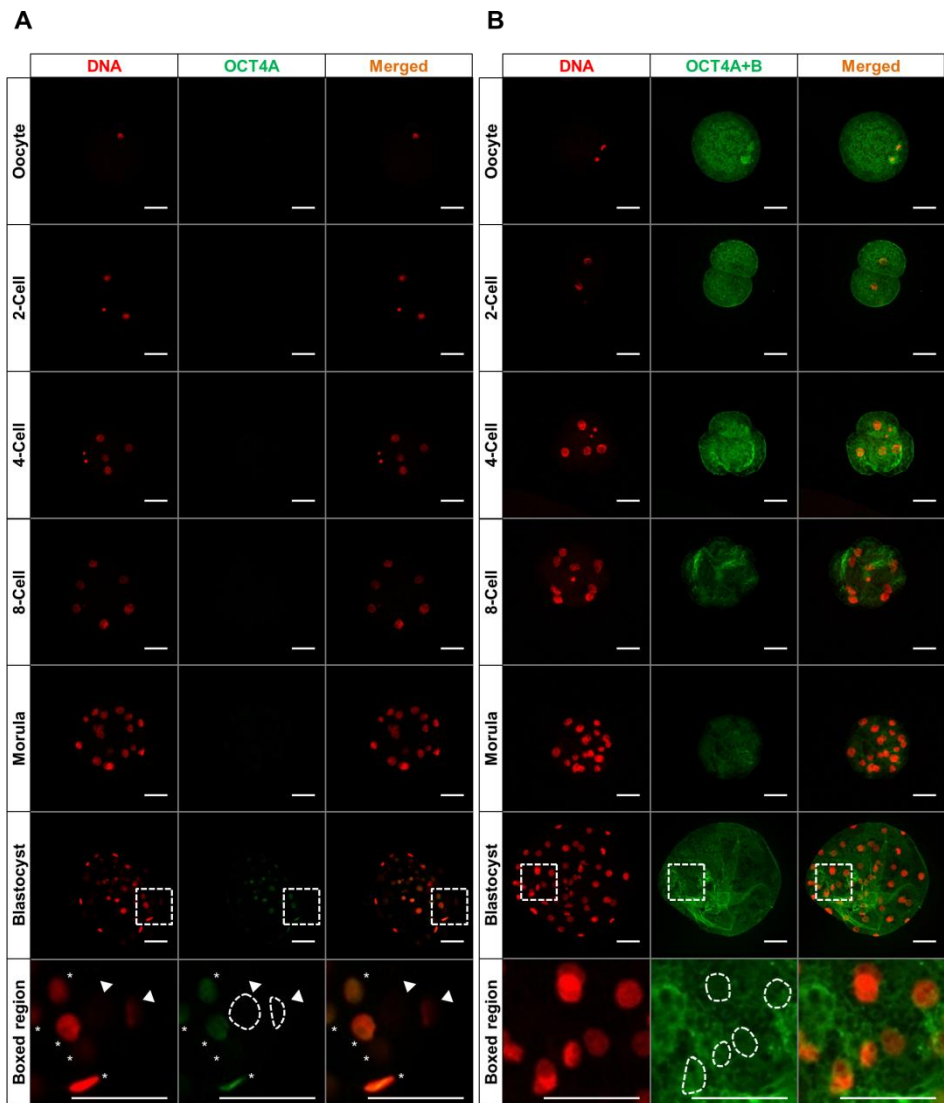


Figure V.9. OCT4 variant protein expression in oocytes and preimplantation embryos. OCT4A (A) and, both OCT4A and OCT4B (B) protein expression was examined in oocytes and developing embryos. DNA and protein targeted by each antibody are represented with red and green, respectively. Asterisks and arrowheads indicate OCT4A-positive and -negative cells in blastocyst, respectively. The area with a dashed line in the magnified region shows the DNA-stained area. Staining with secondary antibodies without primary antibodies was used for a negative control (Figure V.10). Scale bar = 50 μm .

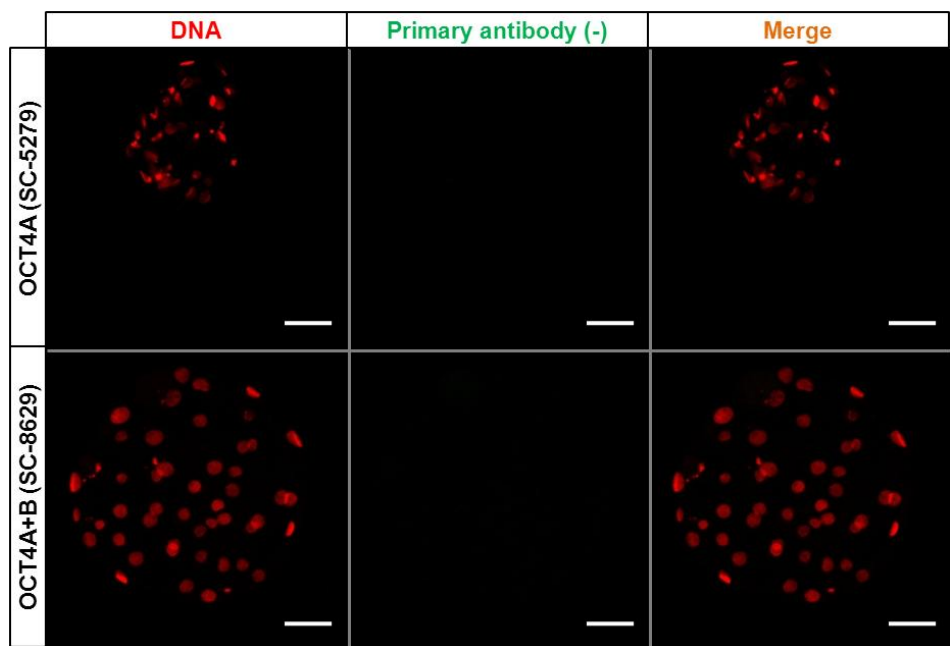


Figure V.10. Immunostained blastocyst without primary antibody. The negative control of each antibody was examined by staining the blastocyst without any primary antibody. Positive signals were not detected in either case. Scale bar = 50 μm .

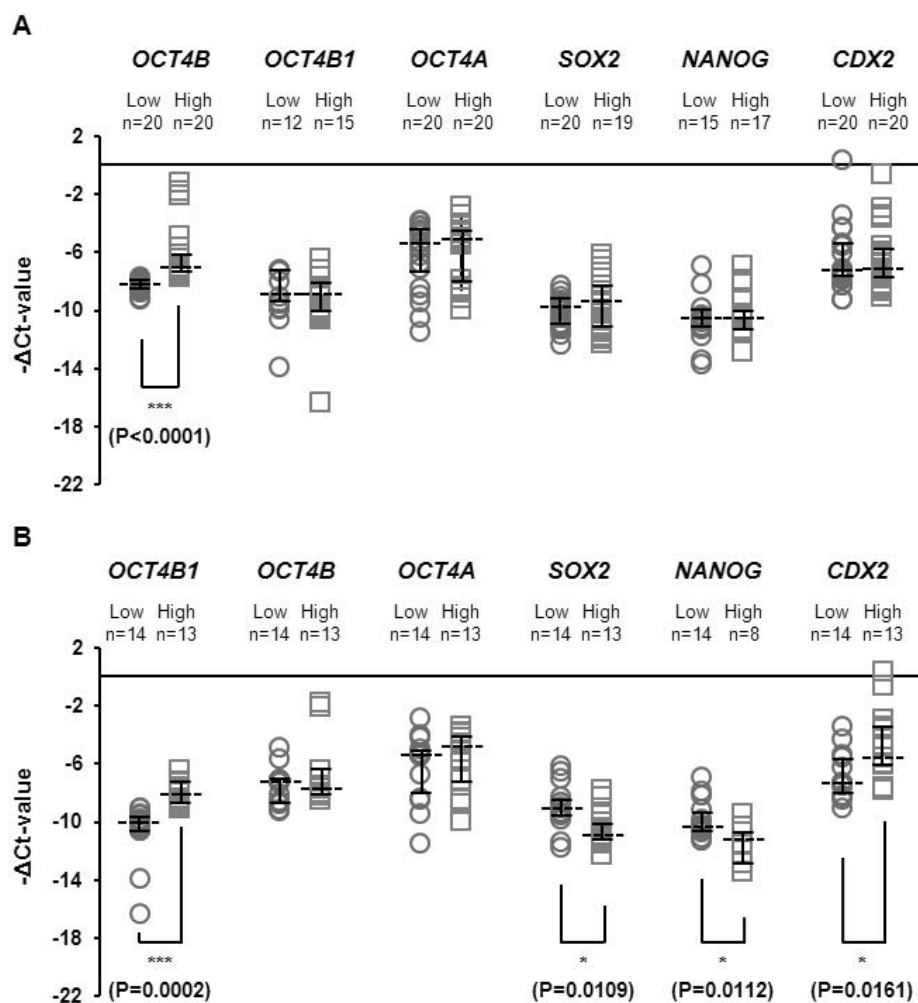


Figure V.11. Pluripotent gene expression comparison of two groups of individual blastocyst classified by expression level of each variant. The aligned dot plot shows the relative quantity of the examined genes from a single blastocyst classified by expression abundance of *OCT4B* (A) and *OCT4B1* (B). The median of *OCT4B* (n=40) and *OCT4B1* (n=27) $-\Delta\text{Ct}$ values (-7.7908 and -9.1157, respectively) were used as standards for grouping individual blastocysts. $-\Delta\text{Ct}$ values of the genes in the blastocyst variant-high group and variant-low group are indicated by circles and squares, respectively. The box plot in each column indicates quartiles. *ACTB* was used as a reference gene and asterisks show significant differences between two groups (*, $P<0.05$ and ***, $P<0.001$).

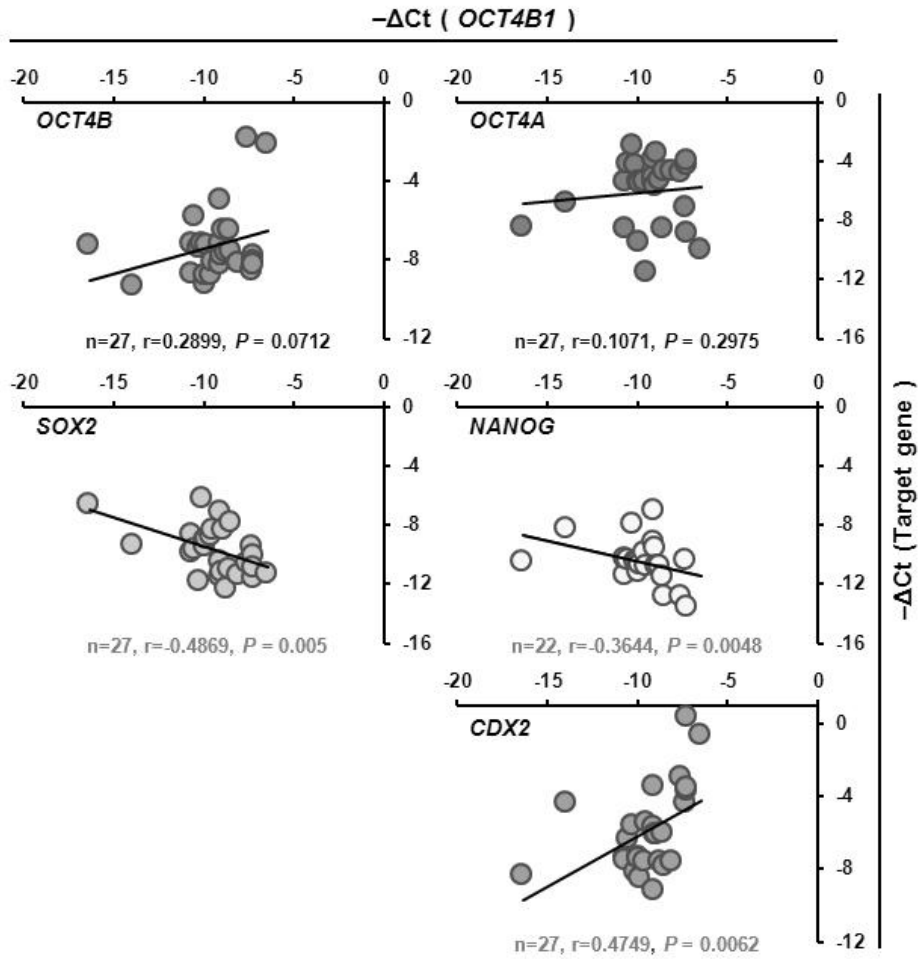


Figure V.12. Correlative expression analysis for *OCT4B1* and pluripotent genes in individual blastocyst. The correlative expression between *OCT4B1* and other variants (*OCT4A*, $n=27$; *OCT4B*, $n=27$) or pluripotent genes (*SOX2*, $n=27$; *NANOG*, $n=22$; *CDX2*, $n=27$) was examined. $P<0.05$ is considered to indicate a significant correlation by the Pearson correlation coefficient. X-axis and y-axis indicate $-\Delta\text{Ct}$ value of *OCT4B1* and examined genes, respectively.

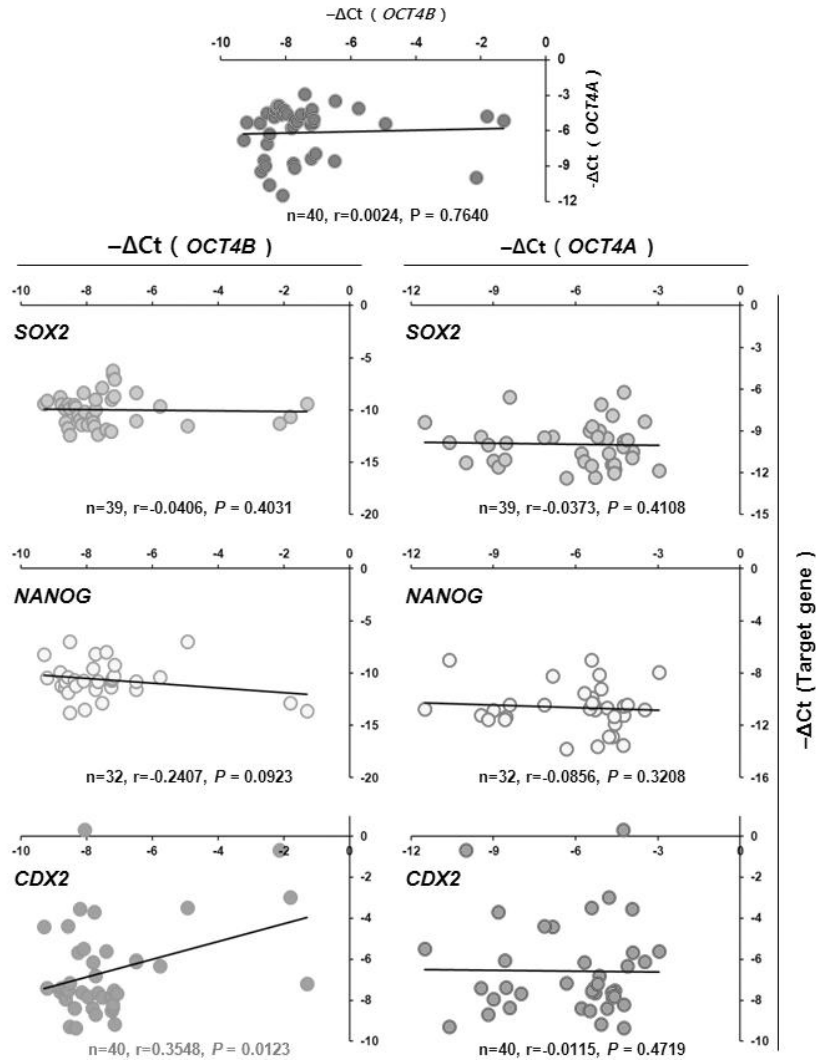


Figure V.13. Correlation analysis between *OCT4* variant and pluripotent genes.

Expression amounts of *OCT4A* and *OCT4B* were correlatively compared with gene expression levels of *SOX2* (n=39), *NANOG* (n=32), and *CDX2* (n=40) in blastocyst. Each circle indicates the expression level of the individual blastocyst. X-axis and y-axis indicate $-\Delta\text{Ct}$ value of *OCT4* variants and analyzed target gene, respectively.

5. Discussion

Previous studies have suggested that *Oct4*, especially *Oct4A*, is the most important gene for maintaining pluripotency in ES cells, acquisition of pluripotency in somatic cells, and normal development in the preimplantation embryo (Nichols *et al.* 1998; Niwa *et al.* 2000; Takahashi and Yamanaka 2006). These studies have led *OCT4A* as an essential gene for analyzing stem cell status and embryo development in various species, including pig. However, compared to the *OCT4A*, other variants have less focused even though several years are passed from the identification of the variant in human and mouse (Takeda *et al.* 1992; Mizuno and Kosaka 2008). The *OCT4* variants have been focused on following two aspects generally, (1) distinct properties compared to the *OCT4A* (Cauffman *et al.* 2006; Lee *et al.* 2006b; Atlasi *et al.* 2008; Guo *et al.* 2012), and (2) inducer making the improper *OCT4A* detection (Wang and Dai 2010). However, these studies were also restricted in human and mouse until recently, and even the presence of *OCT4* variants is still unclear in pig. Therefore, careful insights on the *OCT4* variant are required for the comprehensive understanding of *OCT4* gene as well as accurate analysis of *OCT4A* in pig. So undefined porcine *OCT4* variants, *OCT4B* and *B1*, were confirmed and the expression of each variant was examined independently in developing porcine embryos in this study.

The gene structure of porcine *OCT4B* and *OCT4B1* are conserved

OCT4A sequences and domains among species were highly conserved and the gene's function was suggested as an ancient trait of vertebrates (Tapia *et al.* 2012). The structure of variant mRNA is similar between human and mouse, including POU- and CTD-coding exons with differential NTD-coding exon compared to the A form. However, the sequence of the *OCT4B*-specific exon and NTD were not conserved between human and mouse (Mizuno and Kosaka 2008; Guo *et al.* 2012). In present result, the porcine genome aligned to the human *OCT4B*-specific exon with low homology (Figure V.2) but the transcript structure was similar to that of the previously reported B form of *OCT4* in human (Figure V.3). The other *OCT4* variant, *OCT4B1*, was identified only in human (Atlasi *et al.* 2008). This variant has a cryptic exon (the second intron in *OCT4*) and a stop codon is present on the exon (E2B). Although the lengths and sequences of the *OCT4B1*-specific exons (E2B) were not similar between human and pig (Figure V.2), the variant structure is similar to that of human *OCT4B1* (Figure V.3) and there is a stop codon in the porcine *OCT4B1*-specific exon, as in human (Table V.5). These mean the transcript structure of porcine *OCT4B* and *OCT4B1*, not their sequences, are conserved to those of mouse and human.

Porcine *OCT4B* mRNA generate protein

Transcribed *OCT4B* mRNA has been shown to produce three alternatively translated proteins in mouse and human (Wang *et al.* 2009; Guo *et al.* 2012). These reports also suggested that a few proteins, which are smaller than the protein produced from the first translation-starting site, were produced by an internal ribosomal entry site (IRES). Although an approximately 25 Kda-sized protein, which would be translated by the IRES mechanism as in other species, was detected only (Figure V.7), it is hard to eliminate the possibility that the other sizes of proteins could have been produced alternatively in other tissues which were not examined in this study. Therefore, it is still necessary to examine the presence of alternatively translated proteins by transfecting the *OCT4B* construct in pig cells. Also, the starting codon of the detected protein is unclear because the possible porcine OCT4B ORFs are various compared to mouse and human B-form ORFs. So, confirmation of the exact starting codon is also needed. Meanwhile, results in here demonstrate that the porcine *OCT4B* transcript produces a novel protein, which is different from what was previously reported for OCT4A. The *OCT4B1* mRNA was found to be generated as a truncated protein because of the presence of a stop codon (TGA) in the cryptic exon (Farashahi Yazd *et al.* 2011). Although pig is expected to also produce truncated OCT4B1, as in human, whether or not it does is unclear because of the lack of a proper antibody. Therefore, the presence of porcine OCT4B1 needs to be studied further.

All of *OCT4* variants are expressed in developing embryo

Quantitative analyses of *OCT4* expression in porcine embryonic stage have been carried out in a few previous studies (Lee *et al.* 2006a; Magnani and Cabot 2008) but the expression patterns of the variants were not examined in pig. It was confirmed that the expression patterns of *OCT4B* were similar to that of *OCT4A* but *OCT4B1* expression was variable (Figure V.8A). The proportion of each variant in embryonic stage showed that the variants occupied small proportion (Figure V.8B) as in human embryonic stem cells (Lee *et al.* 2006b). This means *OCT4A* is mainly expressed in developing porcine embryos compared to other two variants, *B* and *B1* form. The protein produced from each variant was also expressed in different stages and located distinctly in developing embryo (Figure V.9). In human, *OCT4B* was produced in the earlier embryonic stages with a different location compared to expression of *OCT4A* (Cauffman *et al.* 2006). As in human, although the exact embryonic stage expressing each variant was different in pig, the results were similar in that *OCT4B* was present in the embryo cytoplasm and was expressed in different embryonic stages compared to what was observed for *OCT4A*. One report showed that the NTD of *OCT4A* is essential for incorporation of the protein into the nucleus (Lee *et al.* 2006b). Porcine *OCT4B* ORFs showed the deletion of the *OCT4A* NTD (Table V.5). This deletion could result in the differential location of *OCT4B* compared to *OCT4A* in pig. These *OCT4* variants expression patterns in embryonic stage demonstrate that they would have the functions in embryo but, the role of the variants could be distinct each other in pig.

***OCT4B* would not be a “pluripotent” marker in pig**

The roles of *OCT4* novel variants are still unclear, in contrast to that of their major form, *OCT4A*. *OCT4B* expression was confirmed in human pancreatic islet (Takeda *et al.* 1992) and mouse eye, especially in retinal pigment epithelial cells (Mizuno and Kosaka 2008). The expression of *OCT4B* in somatic tissue indicates that the variant could have a different function compared to *OCT4A* and would have little relation with pluripotency. In human ES cells, *OCT4B* could not maintain pluripotency in place of *OCT4A* and it failed to bind to the octamer consensus despite the presence of the POU-domain (Lee *et al.* 2006b). Furthermore, one report showed that *Oct4B* failed to reprogram somatic cells into pluripotent stem cells when co-infected with other defined reprogramming factors, *SOX2*, *MYC*, and *KLF4* in mouse (Guo *et al.* 2012). These studies clearly demonstrated that *OCT4B* is not related to pluripotency. In this study, *OCT4B* was detected in all examined somatic tissues (Figure V.7A) and cytoplasmic localization of *OCT4B* (Figure V.9B) showed that the protein would not be a transcription factor. The non-significant correlation with pluripotent gene expression in individual blastocyst (Figures V.11A and 13) also supports the notion that *OCT4B* would not be a pluripotent marker in pig. Although the correlative study suggests restricted information compared to the biological approach, these results demonstrated the possibility that *OCT4B* could be involved in biological functions other than pluripotency in pig, as it is in other species. Some reports showed that *OCT4B* expression was affected by heat and oxidative stress (Wang *et al.* 2009; Guo *et al.*

2012). Therefore, it is possible that porcine *OCT4B* could have a function related with stress-response. These previous reports and expression patterns in tissues and preimplantation embryos indicate that *OCT4B* would not be related to pluripotency as *OCT4A*.

Is *OCT4B1* related to pluripotency in pig?

The other *OCT4* variant, *OCT4B1*, which has been previously identified only in human (Atlasi *et al.* 2008), has been suggested to be involved in pluripotency. Atlasi and his colleagues identified the *OCT4B1* transcript and suggested that *OCT4B1* could have a pluripotency-related function for the following three reasons; 1) *OCT4B1* was highly expressed in ES and EC cells compared to somatic cells; 2) differentiated ES cells lose *OCT4B1* expression; 3) the SSEA3-positive cell population in Shef5 human ES cells strongly expressed *OCT4B1*. Although these results could indicate that *OCT4B1* might be an alternative pluripotent marker (Papamichos *et al.* 2009), the exact biological function of *OCT4B1* is still unclear. In this study, porcine *OCT4B1* was detected in embryonic stage and testis only among the examined tissues (Figures V.3 and V.7A). And correlative expression between *OCT4B1* and pluripotent genes was observed (Figure V.12). Although these results would mean *OCT4B1* might be related to the pluripotency in pigs, it is hard to confirm whether the variant is associated with pluripotency as the results in this study provide only limited evidence via indirect analysis. Some studies tried to

confirm the biological function of *OCT4B1*. Knock-down of *OCT4B1* increased caspase-3 and -7 activity and induced apoptosis in gastric cancer (Asadi *et al.* 2011), and heat-stress increased *OCT4B1* expression (Farashahi Yazd *et al.* 2011). Therefore exact function of the *OCT4B1* and its relation with pluripotency is needed to be studied further.

In this study, porcine *OCT4* variants were identified first. There is a report addressing that discriminative analysis of *OCT4* variants is needed because of their similar structures (Wang and Dai 2010). As in other species, the presence of porcine *OCT4* variants suggests that *OCT4* gene is analyzed carefully by distinguishing other variants also in pigs. Examination of porcine *OCT4* variant expression was carried out in adult tissue, mature oocytes, and developing embryos and it was confirmed that each variant have differential expression patterns. For the comprehensive understanding of porcine *OCT4* genes, *OCT4A* as well as *OCT4B* and *OCT4B1*, their functional studies should be conducted further.

CHAPTER VI

**The effect of *OCT4* on expression of XIC-linked
genes in porcine blastocysts**

1. Abstract

X-chromosome inactivation (XCI) is an epigenetic process for compensating dosages of X-borne genes between male and female eutherians. This process is observed in early eutherian embryo developments with species-specific manner. Until recently, various molecular factors, like non-coding RNAs in X-chromosome inactivation center (XIC) and pluripotent factors, have been suggested to regulate XCI process by activating or repressing *XIST* expression, which is master inducer for XCI. However, insights of the process and its regulators have been restricted in mouse species despite evolutionary diversities on the process and molecular mechanism among the species. Therefore, in here, the relation between one of the most represent pluripotent factors, *OCT4*, which is gate-keeper for maintaining pluripotency and suggested to *XIST* repressor, and three XIC-linked genes, *XIST*, *LOC102165544*, and *RLIM*, was examined in porcine preimplantation embryos. To know their relation, expression levels of *OCT4* and the three genes in blastocysts were correlatively compared. Unexpectedly, expression levels of *OCT4* were positively correlated with *XIST* and *LOC102165544* in female blastocysts. And also, overexpression of exogenous human *OCT4* by lentiviral infection in cleaved embryos generated blastocyst with expressing *XIST* high level. However, the increased *XIST* expression was not observed in blastocysts which were obtained exogenous human *OCT4* in early blastocyst. These results couldn't clearly

demonstrate the roles of *OCT4* on *XIST* expression in porcine blastocyst, but *OCT4* might be involved in molecular networking at earlier stage embryos than blastocysts in pigs.

Key words: *OCT4*, *XIST*, X-chromosome inactivation center, Overexpression, Preimplantation embryos

2. Introduction

Inactivation of X-chromosome in female eutherians is epigenetically essential process to achieve normal embryo development by balancing dosages of X-linked genes between male and female embryos. This process is known to be regulated by X-chromosome inactivation specific transcript (*XIST*). This non-coding RNA (ncRNA) gene has been reported to be a key factor for initiating X-chromosome inactivation (XCI) (Penny *et al.* 1996) and ectopic expression of the gene in cloned mouse embryos revealed differential expression of X-linked genes compared to the fertilized embryos (Inoue *et al.* 2010). This means delicate and accurate regulation of the *XIST* expression is required to success the XCI in embryos. Various studies have revealed that *XIST* expression and initiating XCI is regulated by numerous factors. The genes or enhancers located on the X-chromosome inactivation center (XIC) have been known to regulate *Xist* expression positively (*Rlim*, *Ftx*, and *Jpx*) (Jonkers *et al.* 2009; Tian *et al.* 2010; Chureau *et al.* 2011) or negatively (*Tsix*, *DXpas34*, and *Xite*) (Debrand *et al.* 1999; Lee and Lu 1999; Ogawa and Lee 2003). Add to the factors, pluripotent factors like *Oct4*, *Sox2*, *Nanog*, and *Rex1* were reported as negative regulators for expression of *Xist* and its positive regulator, *Rlim*, and suggested to prevent initiating XCI in mouse ES cells (Navarro *et al.* 2008; Donohoe *et al.* 2009; Navarro *et al.* 2011). These reports showed that the

process for inactivating one of the female X-chromosome is a complex event controlled accurately with numerous regulators in mice. However, unfortunately, the knowledge regarding regulators for *XIST* expression has been restricted in mice until recently and those of other species including pig is still unclear even though the strategy for inducing XCI in early developing embryo is various among the species (Okamoto *et al.* 2011).

In pigs, recent studies identified *XIST* gene (Hwang *et al.* 2013a) and analyzed expression patterns of *XIST* and XIC-linked genes in developing embryos (see the Chapter VI). These results showed that genomic feature of porcine *XIST* and XIC are similar to those of human rather than mouse. And it was suggested that XCI would be initiated but not be achieved chromosome widely in porcine blastocyst. These studies raised the possibility that XCI might be processed differentially between mouse and pig, and the regulators related to the mechanism would be different among the species. However, as the studies represented only restricted knowledge on XCI in pig, and the relation between XCI and their regulators, especially pluripotent markers, are unclear. Therefore, more studies are needed to be carried out for comprehensive understanding of network of the factors in porcine XCI as well as the epigenetic event among the species.

OCT4 is one of the most famous pluripotent genes among the various factors regulating XCI. The gene has been well-known for gate-keeper maintaining pluripotency in inner cell mass (ICM) and its derivate, embryonic stem (ES) cells

in mice (Nichols *et al.* 1998). Interestingly recent report demonstrated that *OCT4* orthologs which are genetically well-conserved gene share the function as a pluripotent marker in vertebrates (Tapia *et al.* 2012). However it is unclear whether the gene also share the functions associated with XCI and blocking *XIST* expression among the species or not. As described above, other regulators, especially ncRNAs in XIC, have been known to have low sequence homology caused by rapid evolution (Romito and Rougeulle 2011) and the mechanism of XCI has been considered to be various among the species (Okamoto *et al.* 2011). Considering the differences on genetic and functional conservation between *OCT4*, and XIC and XIC-linked genes, examining whether *OCT4* is associated with XCI in non-mouse species is important.

So in this study, the relation between *OCT4* and expression of XIC-linked genes is examined in porcine preimplantation embryos. To confirm the relation, expression levels of *OCT4* and XIC-linked genes which were reported to XCI regulators in mouse was compared correlatively. After that, exogenous human *OCT4* was transduced by lentiviral infection into developing embryos. However, unexpectedly, exogenous *OCT4* orthologs led to elevation of endogenous *XIST* expression in porcine female blastocyst in this study.

3. Materials and Methods

Ethics statement

All experiments are conducted with approval of Institutional Animal Care and Use Committees, Seoul National University (SNU-140325-3).

***In vitro* embryo production**

***In vitro* maturation** Prior to generate *in vitro* produced embryos, oocytes were matured in *in vitro* following a previous report (Hwang *et al.* 2013b). Ovaries of pre-pubertal gilt were gifted from Sooam Biotech Research Institute (Seoul, Korea). Cumulus-oocyte-complexes (COCs) were extracted from 3 – 6mm size of follicles and those with multiple layered cumulus cells and granulated cytoplasm was selected. Prepared COCs were washed using TL-Hepes-PVA (Funahashi *et al.* 1997) and cultured in tissue culture medium (TCM-199; Life Technology, Rockville, MD, USA) containing 10% follicular fluid, 10 ng/ml epidermal growth factor (EGF), and µg/ml insulin (Sigma-Aldrich, St. Louis, MO, USA) for 44 hours at 39°C in a 5% CO₂ conditions. Human and equine chorionic gonadotropins (4 IU/ml, Intervet, Boxmeer, Netherlands) were treated for first 22 hours. Forty-four hours later, expanded cumulus cells were detached from the zona pellucida (ZP) by gentle

pipetting with 0.1% of hyaluronidase (Sigma-Aldrich, St. Louis, MO, USA). The oocytes without cumulus cells were assessed to *in vitro* fertilization and parthenogenesis.

***In vitro* fertilization** Fertilized embryos were produced in *in vitro* following a previous report (Hwang *et al.* 2013b). Commercial semen from ducoc breed (DARBI A.I center, Jochiwon, Korea) was assessed and the sperm was washed with Dulbecco's phosphate buffered saline (DPBS; Welgene, Seoul, Korea) supplemented with 0.1% bovine serum albumin (BSA; Sigma-Aldrich, St.Louis, MO, USA) before use. Each twenty to twenty-five of mature oocytes and prepared sperm (1×10^5 cells/ml) were co-incubated on the modified tris-buffered medium (mTBM) (Abeydeera and Day 1997) for 6 hours at 39°C in a 5% CO₂ conditions. After the incubation, the binding sperms were detached from oocytes by gentle pipetting and oocytes were moved to porcine zygote medium 3 (PZM3) (Yoshioka *et al.* 2002) and cultured at 39°C in a 5% CO₂ and 5% O₂ conditions.

Parthenogenesis To generate parthenotes, denuded oocytes were activated by electric pulse (1.0 kV/cm for 60μsec) using BTX Electro-cell Manipulator (BTX, CA, USA) in activation medium (280mM mannitol, 0.01mM CaCl₂, and 0.05mM MgCl₂). The oocytes were transferred to PZM3 containing 2mM 6-dimethylainopurine (6-DMAP; Sigma-Aldrich, St. Louis, MO, USA) and incubated for 4 hours at 39°C in a 5% CO₂ and 5% O₂ conditions. After 4 hours, the oocytes were moved to PZM3 without 6-DMAP, and incubated with same condition to *in*

vitro fertilized embryos.

RNA extraction and Reverse transcription

RNA of individual blastocyst without ZP was extracted by using Dynabeads® mRNA DIRECT™ Kit (Invitrogen, Carlsbad, CA, USA) in accordance with manufacture's instrument. Extracted RNA from blastocysts were reverse-transcribed with High Capacity RNA-to-cDNA™ Kit (Applied Biosystems, Foster City, CA, USA) following the manufacturer's instrument.

Genomic DNA extraction

Genomic DNA (gDNA) from HEK293-LTV cells were extracted by using G-DEX™ IIc Genomic DNA Extraction Kit (iNtRON Bio Technology, Seongnam, Korea) following the manufacturer's instrument. Extracted gDNA was applied for titration of virus.

Quantitative RT-PCR

Quantitative RT-PCR was performed using 0.1 μM of primer sets (Table VI.1) and DyNAmo HS SYBR Green qPCR kit (Thermo Scientific, Rockford, IL, USA)

following the manufacturer's guidance. The reaction was conducted following conditions, one cycle of 50°C for 5 minutes; one cycle of 95°C for 5 minutes; 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. *ACTB* and *RN18S* were used for reference gene.

Table VI.1. Primer pairs used in Real-time PCR

Gene symbol	Primer sequence (5 to 3)	Amplicon size
<i>EGFP</i>	F CTACGGCAAGCTGACCCTGA R TCGCCCTCGAACTTCACCTC	237 bp
<i>hOCT4</i>	F CAGGTGGTGGAGGTGATGGG R CAGAACTCATAACGGCGGGGG	166 bp
<i>OCT4A</i>	F GCTGGAGCCGAACCCCGAGG R CACCTTCCCAAAGAGAACCCCCAAA	150 bp
<i>XIST</i>	F GCTCCAACCAATCTAAAAGGA R ATGCCCCATCTCCACCTAA	131 bp
<i>LOC102165544</i>	F CTAAGATGGCGGCGTTTG R TTGTTTTTCAGGGAATAGAGAGG	135 bp
<i>RLIM</i>	F CCCACCACCGCAAAACTC R CGGCTCACTGCTCTCCAA	159 bp
<i>RN18S</i>	F ACAAATCGCTCCACCAACTAAGA R CGGACACGGACAGGATTGAC	90 bp
<i>ACTB</i>	F GTGGACATCAGGAAGGACCTCTA R ATGATCTTGATCTTCATGGTGCT	131 bp

Lentiviral vector plasmid cloning

Lentiviral vector plasmid coding *enhanced green fluorescent protein (EGFP)* and human *OCT4 (hOCT4)* were prepared. *EGFP* and *hOCT4* insert was obtained by PCR amplification from pLL3.7 (Invitrogen, Carlsbad, CA, USA) and FU-tet-o-*hOCT4* (Addgene, Cambridge, MA, USA) respectively. Each amplicon was produced by using following primer pairs; 5'- ATGGTGAGCAAGGGCGAG-3' (forward) and 5'-CTACTTGTACAGCTCGTCCATGC-3' (reverse) for *EGFP*, and 5'-ATTCGCCACCATGGCGG-3' (forward) and 5'-AATTC(actagt)TCAGTTTGAATGCATG-3' (reverse) for *hOCT4*. Reverse primer for *hOCT4* was designed to have restriction enzyme site of SpeI. Reaction was conducted by using 2x PCR master mix solution (i-MAX II) (iNtRON Bio Technology, Seongnam, Korea) with conditions described as follow; one cycle of 95°C for 7 minutes; 25 cycles of 95°C for 30 seconds, 60°C for 30 seconds, and 72°C for 1 minute; one cycle of 72°C for 10 minutes. Each PCR product was cloned into pGEM®-T Easy vector (Promega, Madison, WI, USA), and transformed into E.coli (Novagen, Madison, WI, USA). Plasmid was extracted using DNA-spin™ Plasmid DNA Purification Kit (iNtRON Bio Technology, Seongnam, Korea) and sequenced using ABI PRISM 3730 DNA Analyzer (Applied Biosystems, Foster, CA, USA) to select plasmid harboring insert with correct sequences. Transgene in T-vector was obtained by enzyme digestion and passed to *LIN28*-removed linearized pSIN-EF2-LIN28-PUR plasmid (Addgene, Cambridge, MA, USA) by using T4 DNA ligase (New England BioLabs Inc, Beverly, MA,

USA) and transformed to NEB Stable Competent E.coli (New England BioLabs Inc, Beverly, MA, USA) following the manufacturer's instrument. Plasmid was extracted following same method mentioned above and insertion of transgene was confirmed by digesting plasmid by restriction enzyme.

Lentivirus production

Lentivirus was produced following the previous report (Nagano *et al.* 2002) with modification. Briefly, HEK293 LTV cells (Cell Biolabs, USA) were used for packaging the lentivirus and cultured following the manufacturer's instrument. Chloroquine (25 μ M) was treated to the cells three hours previous to transduction of lentiviral vector plasmid. Four plasmids, self-inactivating lentiviral vector plasmid harboring transgenes (pSIN-EF2-(*EGFP* or *hOCT4*)-PURO), packaging plasmids (pLP1 and pLP2; Invitrogen, Carlsbad, CA), and envelop plasmid (pLP/VSVG; Invitrogen, Carlsbad, CA), were transduced to the chloroquine-treated cells by calcium phosphate precipitation method. Culture supplement was harvested three times every 24 hours after plasmids transduction and preserved at the 4°C. Stored culture medium was filtered using 0.45- μ m pore filters (Sartorius Stedim Biotech, Goettingen, Germany) and 5x polyethylene glycol (PEG) solution (2.5 M of NaCl and 40% PEG; Sigma-Aldrich, St. Louis, MO, USA) was added to filtered solution. The mixture was centrifuged at 10,000xg for one hour and the virus pellets were dissolved to DPBS containing 1 % BSA. Diluted virus was

stored at the - 80°C until use.

Lentivirus titration

Virus titration was performed by calculating the integrated copy numbers in HEK293-LTV cells inoculated by prepared virus.

Copy number calculation To calculate copy numbers of integrated vectors and genome, linear-regression plot was prepared following previous report (Hwang *et al.* 2014). Intron region of human *GAPDH* (*hGAPDH*) and internal ribosome entry site (IRES) in lentiviral vector plasmid were targeted to measure the integrated copy in one cell. PCR product of target region in human genome (*hGAPDH*) and in vector (*pSIN-IRES*) was amplified by using human gDNA extracted from HEK293 LTV cells and pSIN-EF2-LIN28-PUR plasmid (Addgene, Cambridge, MA), respectively, with following primers; 5'-CTCTCTCCCATCCCTTCTCC-3' (forward) and 5'- CCCACCCCTTCTCTAAGTCC-3' (reverse) for *hGAPDH*; 5'-AGGTCTGTTGAATGTCGTGAA-3' (forward), and 5'-CCCCTTGTTGAATACGCTTG-3' (reverse) for *pSIN-IRES*. The reaction was performed using 2x PCR master mix solution (i-MAX II) (iNtRON Bio Technology, Seongnam, Korea) with following conditions; one cycle of 95°C for 7 minutes; 40 cycles of 95°C for 30 seconds, 60°C for 20 seconds, and 72°C for 30 seconds; one cycle of 72°C for 10 minutes. Each product was run on the 1% agarose-gel and

extracted using MEGAquick-spin™ Total Fragment DNA Purification Kit (iNtRON Bio Technology, Seongnam, Korea). Extracted amplicons were diluted 10-fold serially (10^{-5} to 10^{-8}) and assessed to quantitative PCR to obtain Ct-value of each aliquot. Log of copy number and its Ct value was plotted (Figure VI.1A and Table VI.2). The plot was used for calculating copy number of vectors in one cell. Copy numbers of each target region in 1 µl of gDNA sample are calculated as follow;

Copy number of 1 µl sample gDNA

$$= \text{Copy number of 1 } \mu\text{l PCR product} \times 10^{\frac{(Ct-Y_Intercept)}{Slope}}$$

Copy number of integrated vectors in one cell was calculated as follow

Number of integrated vector in one cell

$$= \frac{\text{Copy number of pSIN – IRES}}{2 \times \text{Copy number of hGAPDH}}$$

Titration of viral stock Stocks of lenti-virus having *EGFP* and *hOCT4* expression vector (*EGFP*-lentivirus and *hOCT4*-lentivirus, respectively) were serially diluted 10-fold and transduced to HEK293-LTV cells for one day. The cells inoculated with each concentration of viral stocks were harvested and gDNA was extracted

from the each sample. Extracted gDNAs were assessed to quantitative PCR and Ct values of each sample were obtained. Copy number of integrated vectors in one cell was calculated and, log of copy number was plotted to the log of dilution rate of virus stock (Figure VI.1B). The following equation of the linear regression curve was used for calculating integration rate of each virus stock and matching the

$$\textbf{Infection rate of stock} = 10^{(\textit{Slope} \times \log(\textit{dilution rate}) + \textit{Y_intercept})}$$

Lentivirus transduction into embryos

Lentivirus was co-cultured with one-cell embryos with slit ZP or cleaved embryos (day 3) and early blastocysts (day 5) without ZP to make possible cytoplasm exposed to virus (Figure VI.2). Both lentiviral stocks were diluted to same concentration with PZM3 and treated to embryos for one day. One cell embryos with Slit ZP (SZ) were generated by cutting ZP of mature oocytes by manipulation followed by electronic pulse. After treating 6-DMAP, embryos were exposed to viruses. Viral infected cleaved embryos and early blastocyst was produced as follow. ZP of each stage of embryos were removed by Tyroid's Acid (Sigma-aldrich, St.Louis, MO, USA) and co-cultured with lentivirus. These embryos, SZ and ZP-removed at embryonic day 3 (ZR3) and day 5 (ZR5) were treated with lentivirus for one day followed by washing four times with DPBS

containing 0.4% BSA. Washed embryos were transferred to PZM3 and blastocysts (day 7 from activation) were assessed to mRNA extraction. Virus integration in blastocyst was confirmed by observing green fluorescent (*EGFP*-lentivirus inoculation) or confirming expression of puromycin resistance gene (*PURO*) using PCR (*EGFP*- and *hOCT4*-lentivirus inoculation) with following primers; 5'-3' (forward) and 5'-3' (reverse). The reaction was carried out using 2x PCR master mix solution (i-MAX II) (iNtRON Bio Technology, Seongnam, Korea) with following conditions; One cycle of 95°C for 7 minutes; 40 cycles of 95°C for 15 seconds, 60°C for 20 seconds, and 72°C for 30 seconds; one cycle of 72°C for 10 minutes.

Statistical analysis

Statistical analysis was carried out with the Graphpad Prism statistical program (Graphpad Software, San Diego, CA, USA). Comparisons of developmental competence among the embryos and expression levels of blastocysts inoculated with *hOCT4*- or *EGFP*-lentivirus were performed using Student *t*-test. Analysis of correlative expression between endogenous *OCT4* and XIC-linked genes in individual blastocyst was conducted by Pearson correlation coefficient. All data were exhibited as mean \pm standard error mean (S.E.M) and $P < 0.05$ is considered to be significantly different.

Table VI.2. *pSIN-IRES* and *hGAPDH* information for linear-regression plot

Target gene	Concentration (PCR product, ng/ul)	Molecular weight (g/1M)	Weight / 1 Copy (ng) ^a	Copy Number (1ul Product)	Amplification plot			
					Slope	Y-intercept	Amplification ^c efficiency	R ²
<i>pSIN-IRES</i>	2.09	147102.06	2.44269E-10	8.5562E+09	-3.2023	-2.0027	2.052464017	0.991
<i>hGAPDH</i>	2.00	120532.85	2.00149E-10	9.9925E+09	-3.3897	-2.7176	1.972474321	0.994

^aOne copy of PCR product (ng) = (Molecular weight of PCR product)/(Avogadro number) × 10⁹, Avogadro number = 6.0221413 × 10²³

^bCopy number (1 µl of PCR product) = (PCR product concentration, ng/1ul) / (one copy of PCR product, ng)

^cAmplification efficiency = 10^(-1/slope)

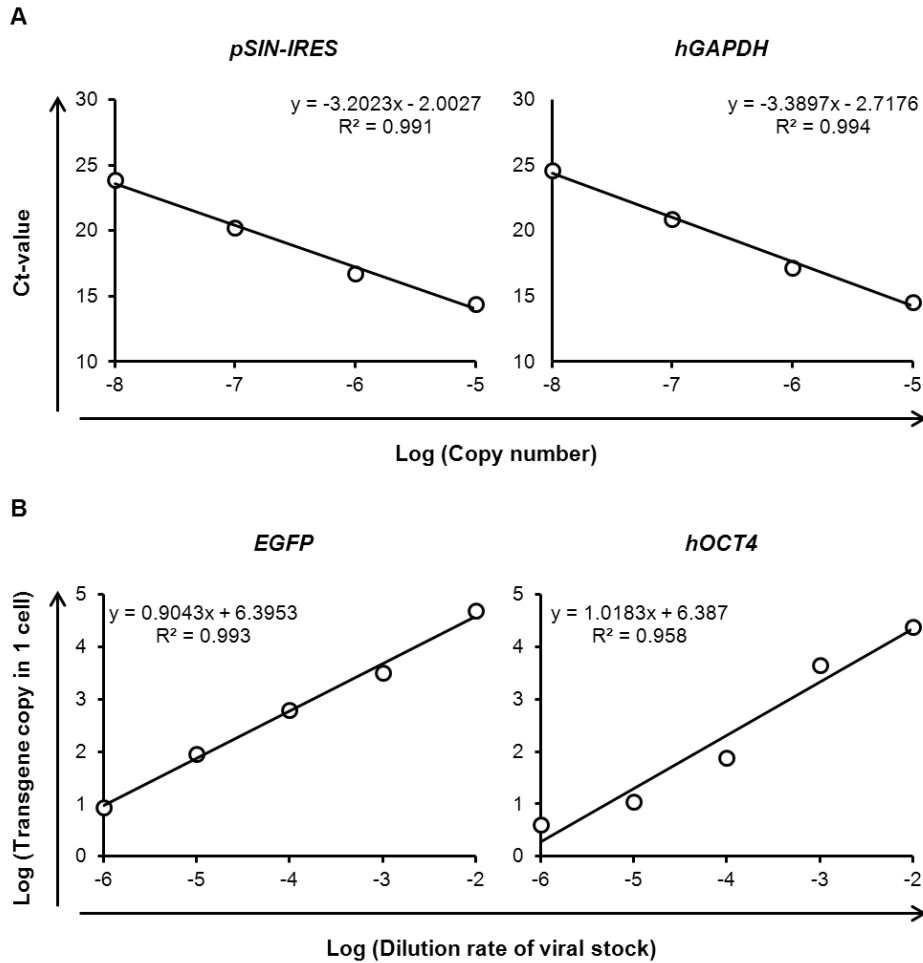


Figure VI.1. Linear-regression plot for calculating copy number of transgene and integration efficiency of lentivirus stock. (A) Plot for calculating copy number of host genome and transgene. *pSIN-IRES* and *hGAPDH* was used for measuring copy number of transgene and host genome (HEK293 LTV), respectively. (B) Standard curve for calculating integration rate of each lentivirus to HEK293 LTV. The plots indicate the number of transgene copy in one cell through the treated virus stock concentration.

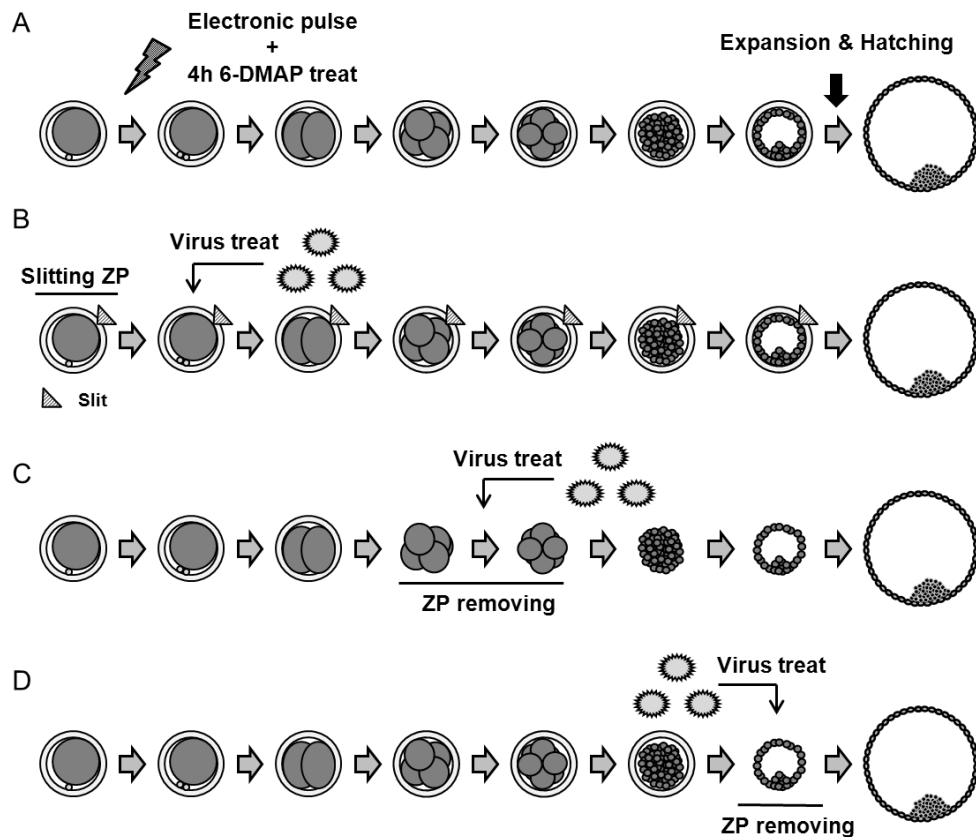


Figure VI.2. Strategies for lentivirus transduction in developing embryos. To infect the lentivirus harboring transgenes into porcine developing parthenotes, various strategies were applied. Manipulating the zona pellucida (ZP) was carried out to expose cytoplasm of parthenotes which were generated by electronic pulse and 4 hours treatment of 6-DMAP (A) directly to the medium containing lentivirus. To make a slit, ZP of mature oocyte was cut with glass pipet by manipulation (B). To inoculate the virus at the cleaved embryos (C) and early blastocyst (D), ZP of each stage embryo was removed and virus was infected for one day.

4. Results

Correlative expression analysis of OCT4 and XIC-linked genes in male and female blastocyst

To know the relations between *OCT4* and XIC-linked genes, their expression levels in male and female blastocyst was correlatively compared. *XIST*, and *RLIM*, which were reported to the genes regulated by *Oct4* (*Xist* and *Rlim*) (Navarro *et al.* 2008; Navarro *et al.* 2011), and uncharacterized ncRNA, *LOC102165544*, which was considered to be an ortholog of *Jpx*, another *Xist* activator in mouse (Tian *et al.* 2010), in pigs (See the Chapter IV) were selected to be compared correlatively with endogenous *OCT4* expressions. Sexes of fertilized blastocysts were classified by *XIST* expression as reported previously and expression levels of *OCT4* were not different in male and female blastocysts (Figure VI.3A). Correlative comparison of the expression levels between *OCT4* and XIC-linked genes showed that *OCT4* expression is quantitatively correlated with *XIST* and *LOC102165544* in fertilized female blastocysts, and the similar result was also observed in parthenogenic blastocyst (Figure VI.3B). Contrary to female blastocyst, any expressions levels of XIC-linked genes were not significantly correlated with *OCT4* expression in fertilized male blastocyst. The results raised the possibility *OCT4* expression would be related to XIC-linked genes in porcine female blastocyst.

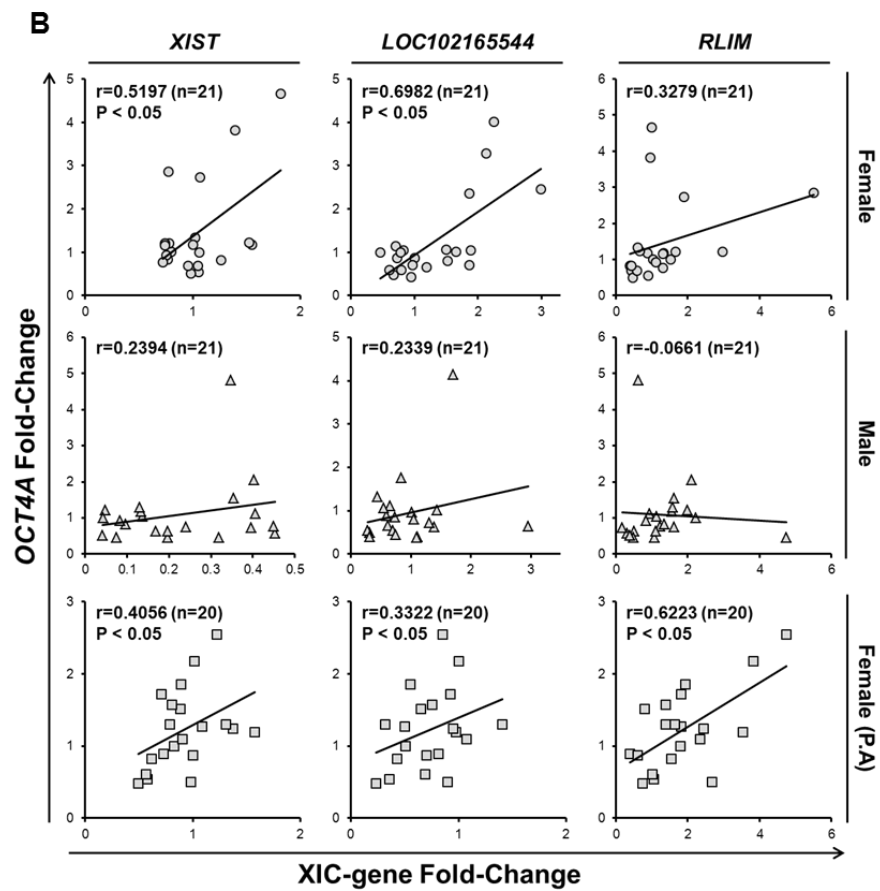
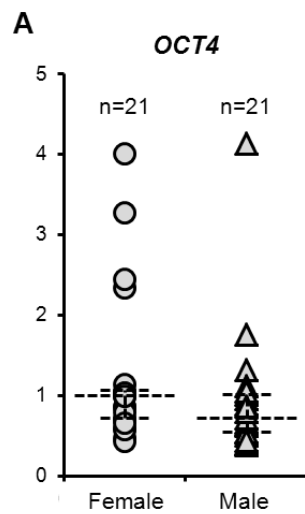


Figure VI.3. Correlative analysis of *OCT4* expression levels with those of XIC-linked genes in blastocyst. (A) *OCT4* expression level comparison between male and female fertilized blastocyst. Fertilized blastocysts were sexed by expression levels of *XIST* as previously reported (Park *et al.* 2011). Each circle and triangle indicates *OCT4* expression levels of individual female and male blastocyst, respectively. Box plot in each column represents quartile of expression level distribution. *ACTB* and *RN18S* were used for reference genes. Median of *OCT4* expression level in female was set to one-fold. (B) Correlative comparison of expression levels between *OCT4* and XIC-linked genes. *ACTB* and *RN18S* were used for reference gene. Medians of each gene expression levels in fertilized female blastocyst were set to one-fold. $P < 0.05$ was considered to be significant correlation.

Slitting or removing ZP don't affect to development of parthenote in *in vitro*

Prior to inoculating lentivirus to embryo, it was confirmed whether manipulating or removing ZP would affect to development of preimplantation embryos. Developmental competence and cell numbers of blastocyst with SZ didn't showed significant difference compared to the blastocyst with intact ZP (Table VI.3). And formation rates of blastocyst from ZR3-embryos were not significantly different to that of cleaved embryos at day 3 (Table VI.4). These results indicate that manipulating or removing ZP wouldn't affect to developmental competence of embryos (Figure VI.4).

Lentivirus transduction in preimplantation embryo

ZP prevents insertion of external harmful factor like integration of virus to growing oocytes and developing embryos by acting as a barrier (Van Soom *et al.* 2010). Previous study reported that the exposing oocyte or embryo cytoplasm to external environment by drilling ZP using laser made possible to viral infection into bovine oocytes (Ewerling *et al.* 2006). Therefore, *EGFP*-lentivirus was treated to zygote with SZ, ZR3-, and ZR5-embryos by incubating together to determine efficient viral infection methods (Figure VI.2). Any approaches for treating the lentivirus didn't showed any effects on embryo development in this study (Table VI.5 and VI. 6). However, successful viral infections and *EGFP* expressions were confirmed

Table VI.3. Effect of slit ZP on parthenotes development

	No. of oocytes	Cleaved embryo (%)		Blastocyst (%)		Cell No. of Blastocyst (n=6)
Slit ZP*	139	110	(79.85 ± 4.89)	64	(46.64 ± 3.40)	43.83±5.59
Control	158	128	(82.46 ± 5.48)	71	(45.35 ± 2.89)	35.17±4.03

*Slit on ZP was produced by cutting ZP of mature oocytes.

Table VI.4. Effect of ZP-removal on cleaved parthenotes development

	No. of embryos*	Blastocyst (%)		Cell No. of Blastocyst (n=7)
ZP-removed	57	19	(33.69 ± 2.66)	36.14 ± 5.75
Control	53	22	(42.05 ± 6.13)	40.14 ± 6.21

*Cleaved parthenotes (day 3) were used.

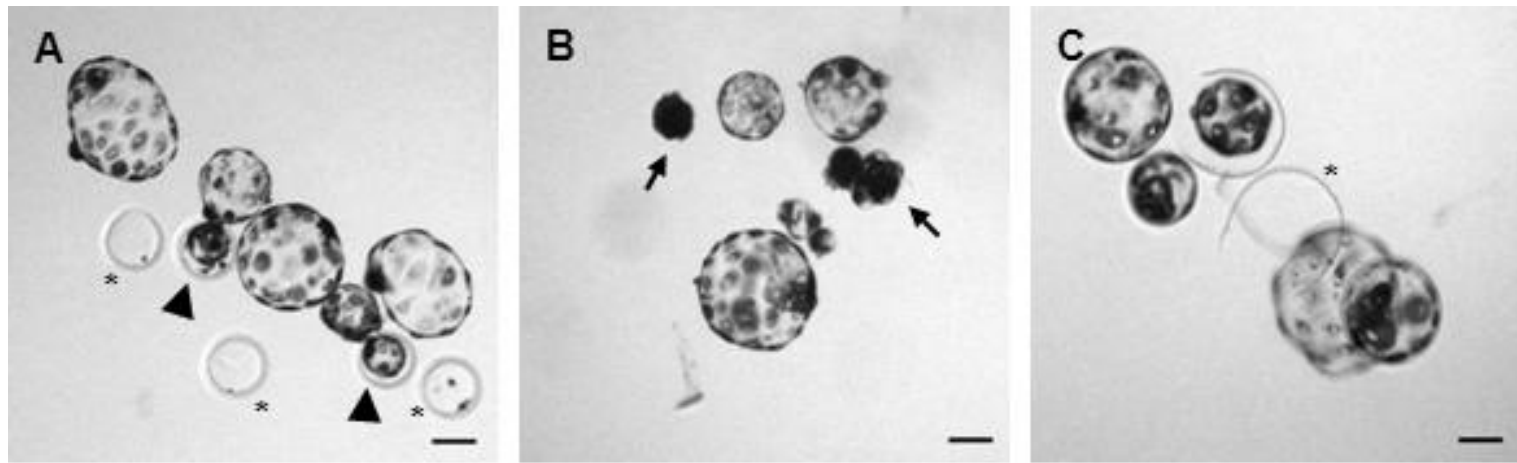


Figure VI.4. Generation of parthenogenic blastocyst with slit or removed ZP. Parthenogenic blastocysts generated from oocytes with slit ZP (A) or cleaved parthenotes of which ZP was removed (B) were exhibited. Oocytes with intact ZP were used for control (C). Asterisks and arrowheads indicate empty ZP and blastocyst showing assisted hatching, respectively. Arrows in panel (B) mean ZP-removed embryos which were developmentally arrested and failed to form blastocysts. Scale bars = 100 μ m.

Table VI.5. Developmental competence of embryos with slit ZP (SZ) cultured with lentivirus					
	No. of oocytes	Cleaved embryo (%)		Blastocyst (%)	
<i>EGFP</i> *	104	78	(75.35 ± 6.95)	32	(30.82 ± 3.51)
Control	112	88	(78.49 ± 2.85)	43	(37.83 ± 7.42)

*Concentrations of virus stock = 2 x 10⁶ virus particles/one cell of HEK293 LTV.

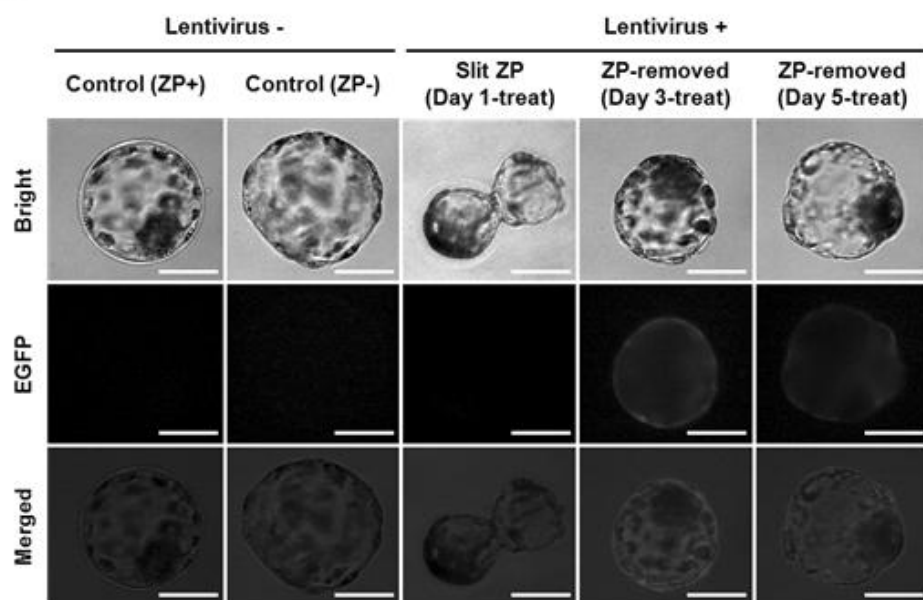
**Parthenotes with slit ZP was used

Table VI. 6. Developmental competence of ZP-removed at embryonic day 3 (ZR3) embryos cultured with lentivirus		
	No. of embryos**	Blastocyst (%)
<i>EGFP</i> *	43	19 (43.81 ± 8.10)
Control	57	19 (33.69 ± 2.66)

*Concentrations of virus stock = 2 x 10⁶ virus particles/one cell of HEK293 LTV

**ZP-removed cleaved parthenotes were used.

A



B



Figure VI.5. Validation of lentiviral infection into porcine embryos. (A) *EGFP* expression in blastocyst infected with *EGFP*-lentivirus at differential embryonic stage. Successful viral infection was observed in blastocysts which were treated with *EGFP*-lentivirus at ZP-removed embryonic (ZR3- and ZR5-embryos), but blastocysts with SZ were failed to be inoculated. Non-viral treated blastocysts with or without ZP were used for controls. Scale bars = 100 μ m. (B) Expression of transgene, *PURO* (203 bp), was examined in blastocysts which had been treated with lentivirus at each embryonic stage. Expression of exogenous gene was not observed in control (line 1-3) and lentivirus co-cultured blastocysts which have SZ (line 4-6). However, blastocysts from ZP-removed cleaved embryo (ZR3-embryos, line 7-9) and early blastocyst (ZR5-embryos, line 10-12) showed successful viral infections and transgene expressions. Blastocyst with intact ZP was used for control and ‘M’ indicates 50 bp-DNA size markers.

when cytoplasm of embryos were thoroughly exposed to virus particles (ZR3- and ZR5-embryos) and expression of transgene in blastocysts with SZ were not observed (Figure VI.5). Therefore, ZP-removed embryos were used for viral infection in this study.

Effect of *hOCT4* overexpression on XIC-linked gene expression in blastocyst

To know the functions of *OCT4* in expressions of X-linked genes in female blastocysts, ZR3-embryos were infected with *hOCT4*-lentivirus (Figure VI.6). Interestingly, blastocysts developed from ZR3-embryos infected with *hOCT4*-lentivirus showed reduced blastocyst formation compared to controls, which was infected with *EGFP*-lentivirus (Figure VI.6A). Elevation of endogenous *OCT4* expression was observed in *hOCT4*-expressing blastocysts. Unexpectedly, expression levels of *XIST* in blastocyst with exogenous *hOCT4* were elevated about 5-fold. However, differential results were observed when the virus was infected into early blastocyst (ZR5) (Figure VI.7). Viral infection into ZR5-embryos didn't influence to the survival rates of blastocysts and expression levels of *XIST* in developed female blastocysts (Figure VI.7) contrary to the blastocysts which were infected with the virus at early stage of embryos. Expression levels of other X-linked genes, *LOC102165544* and *RLIM*, were not changed in blastocysts infected with *hOCT4*-lentivirus on ZR3- and ZR5-embryos, both (Figures VI.6C and VI.7C).

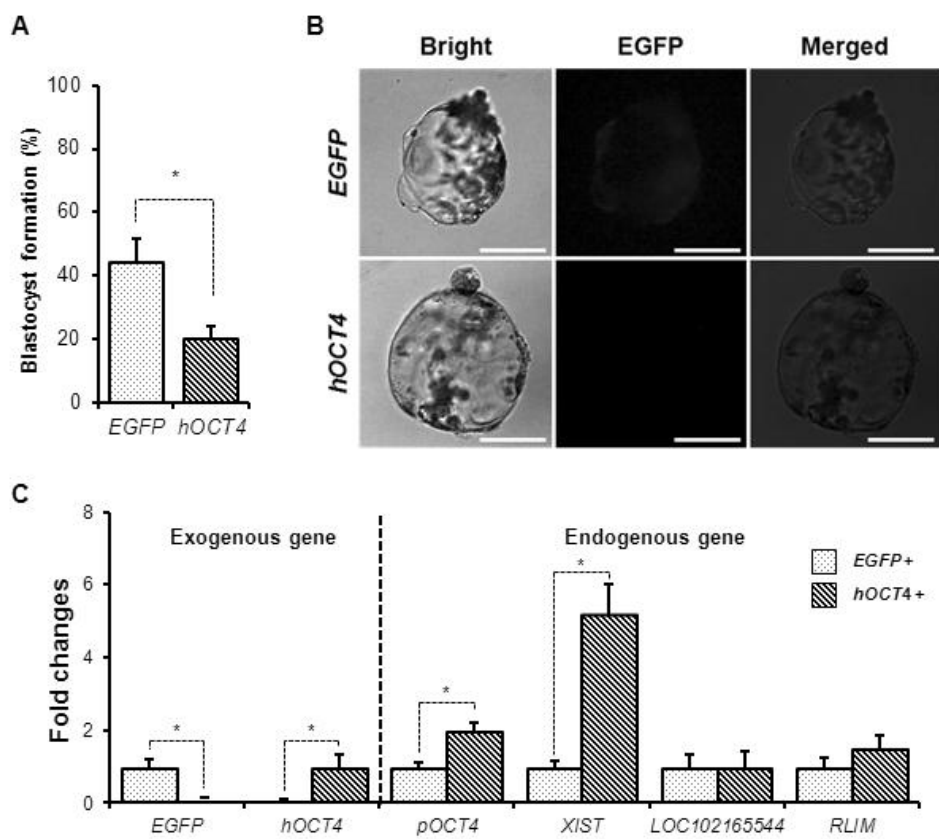


Figure VI.6. Effect of exogenous human *OCT4* on parthenogenic blastocyst infected at cleaved embryos. (A) Blastocyst formation rates of *hOCT4*-expressing cleaved embryos. The percentage of embryos which were developed to blastocysts after viral infection was represented. Blastocysts infected with *EGFP*-lentivirus were used for controls. Experiment was three times replicated. Asterisk indicate significant difference ($P<0.05$). (B) Exogenous *EGFP* expression in blastocyst. Blastocysts which were infected *EGFP*-lentivirus at their cleaved stage were expressed *EGFP* specifically. Scale bars = 100 μm . (C) Relative comparison of expression levels between blastocysts expressing *EGFP* and *hOCT4*. Median of each gene expression levels in *EGFP* expressing blastocyst ($n=7$) was set to one fold. In case of *hOCT4* expression, its median value in *hOCT4* infected blastocyst ($n=7$) was set to one fold. Asterisks indicate significant difference ($P<0.05$).

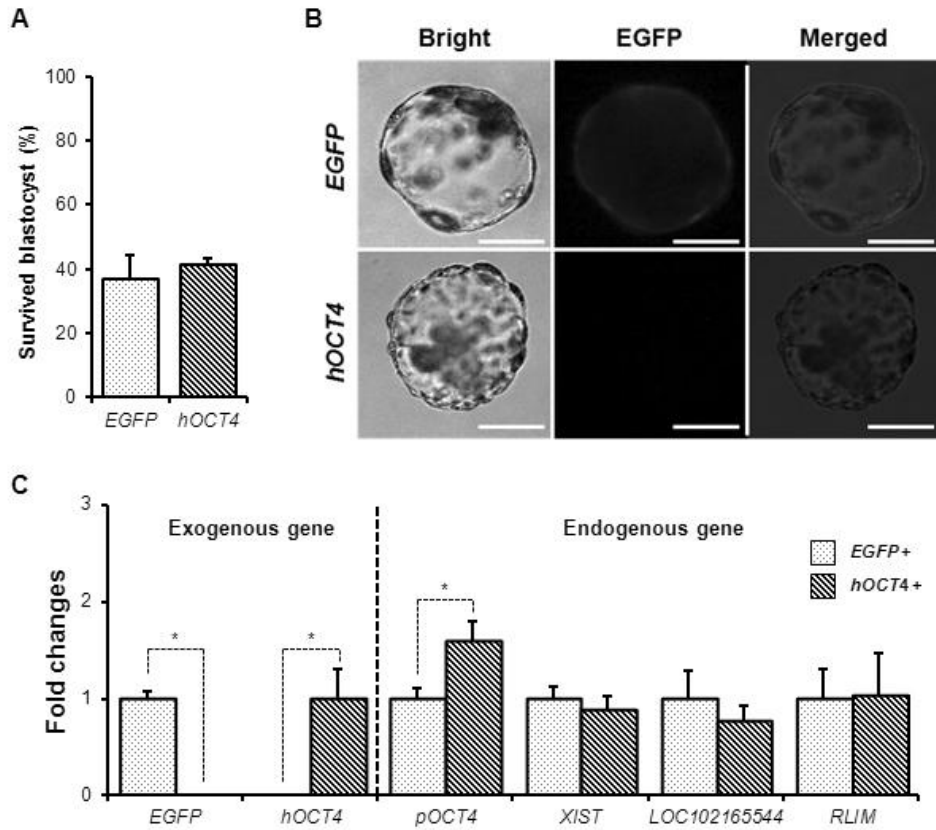


Figure VI.7. Effect of exogenous human *OCT4* on parthenogenic blastocyst infected at early blastocyst. (A) Survival rate of blastocyst infected with *hOCT4*-lentivirus. Survival rate of blastocysts after viral infection at early blastocyst was represented. That of blastocysts infected with *EGFP*-lentivirus was used for control. Experiment was three times replicated. (B) Exogenous *EGFP* expression in blastocyst. Scale bars = 100 μ m. (C) Relative comparison of expression levels between blastocysts expressing *EGFP* (n=5) and *hOCT4* (n=5). Details are same to the legends in Figure VI.6C.

5. Discussion

XCI is a chromosome-wide gene silencing process occurred in X-chromosome of early developing embryos to equalize the expression levels of X-linked genes between male and female individuals. Although numerous studies were conducted to know the mechanisms of this complex epigenetic event and associated regulators using mouse models, their insights in non-mouse species is still rare. Comparative genomic studies showed that XIC-linked ncRNAs regulating XCI have been evolved rapidly and, the orthologs were less conserved and genetically diverse among the mammals (Hore *et al.* 2007; Romito and Rougeulle 2011). Also, one study revealed that the processes for XCI and its initiation timing in developing embryos are variable among the species (Okamoto *et al.* 2011). These studies demonstrate the requirement of species-specific researches on XCI in non-mouse species. So in this study, the relation between *OCT4* and three XIC-linked genes of which orthologs were reported to induce XCI in mouse has been examined to know that function of this represent pluripotent gene on XCI is conserved in pigs.

In mice, *Oct4* has been known to a master factor for maintaining pluripotency in ICM and ES cells (Nichols *et al.* 1998; Niwa *et al.* 2000). One report suggested that the function of this gene is conserved by showing *POU2*, which is an ancestral gene of *OCT4* in tetrapod, could reprogram somatic cells and induced pluripotent

stem cells in humans and mice (Tapia *et al.* 2012). However, although numerous studies were revealed the importance of the genes in pluripotency, a few studies highlighted the roles of *OCT4* in XCI (Donohoe *et al.* 2007; Navarro *et al.* 2008). First report observing the relation between pluripotent factor and *Xist* expression suggested that major pluripotent factors, including *Oct4*, would negatively related with *Xist* by binding to *Xist* intron 1 in mouse ES cells (Navarro *et al.* 2008). They observed the dissociation of the factors from the first intronic region in *Xist* locus in differentiating ES cells and rapid elevation of *Xist* transcript (within 24 hours) compared to reduction of other pluripotent factors which are naturally decreased following the ES cells differentiation (48 to 96 hours) induced by *Oct4* deletion. Another study performed by Donohoe *et al.* observed downregulation of *Oct4* induced elevation of the ES cell population showing ectopic *Xist* expression (Donohoe *et al.* 2009). By observing the interaction between Ctf and OCT4, and its localization on *Tsix* promoter, they suggested that *Oct4* also have *Tsix*-mediated *Xist* repressing roles in mouse ES cells. Even though the two reports suggested differential mechanism of *Xist* repression by *Oct4* in mouse ES cells, it looks clear that that *Oct4* would negatively acts for *Xist* expression and its subsequent result, XIC, at least in mice.

However, the results in here showed opposite tendency compared to the studies in mice. Unexpectedly, correlative analysis of expression levels (Figure VI.3) and exogenous *OCT4* inducing in cleaved embryos (Figure VI.6) showed that *XIST* expression was positively influenced to *OCT4* expression in pigs. Although the

reason for the results is unknown and it is hard to exclude the possibility that opposite function of *OCT4* to *XIST* expression in pig blastocyst, this unexpected elevation of *XIST* by overexpression of exogenous *OCT4* ortholog might be resulted from broken balances of pluripotent molecular networking during early embryo development. Indeed, despite the importance of the *Oct4* in maintenance of pluripotency, overdose of the gene induced differentiation into primitive endoderm and mesodermal cell lineage in mouse ES cells (Niwa *et al.* 2000). In here, expression levels of exogenous *hOCT4* in blastocysts developed from ZR3-embryos were about ten-fold high than control and this might accelerate the differentiation of cells during early embryo development. This upregulated *XIST* expression would induce malformation of blastocyst (Figure VI.6A) because *XIST* expression closely linked to the abnormal embryo development in various species (Wrenzycki *et al.* 2002; Inoue *et al.* 2010; Park *et al.* 2012). However, the elevation of *XIST* expression was not observed in the blastocyst which obtained exogenous *OCT4* at their early blastocyst stage (Figure VI.7). As shown in previous study (Chapter IV), explosive increase of *XIST* expression was observed after morulae. This might indicate starting of XCI after morular stage and the status of cells in early blastocysts might start the segregation into specific cell lineages because XCI and pluripotency is tightly linked in embryonic stage and ES cells (Nichols and Smith 2009; Ohhata and Wutz 2013). Therefore, even though almost fifty-fold increased expression of exogenous factors was observed, this overdose might not affect to the molecular networking of blastocysts in here. However, as exogenous

gene was overexpressed during short period in blastocysts from the ZR5-embryos (2 days) compared to the blastocysts taken *hOCT4* at ZR3-embryos (4 days), this rather short period for expressing *hOCT4* would not affect to *XIST* expression in ZR5-blastocyst.

Expression levels of XIC-linked genes examined in this study didn't show any changes following induction of *hOCT4* (Figures VI.6 and VI.7). Even though *Rlim* was suggested to be regulated by *Oct4* negatively in mouse ES cells (Navarro *et al.* 2011), expression of its ortholog in pigs was not influenced by the overexpression of *OCT4*. This might be resulted from the differential *OCT4* roles to *Rlim* regulation in pigs. One curious thing is why the expression level of *LOC102165544* and *RLIM* was not changed even *XIST* expression levels were elevated in blastocyst. Previous study in Chapter IV showed elevation and reduction of *LOC102165544* and *RLIM* expression, respectively, following the *XIST* increase. Because the *Rlim* mRNA was diminished at the blastocyst stage compared to morula in porcine preimplantation embryos, repression of the gene would be already achieved before increased *XIST* RNA influenced the *RLIM* expression.

In this study, effect of the *OCT4* overexpression to XIC-linked gene expressions was examined during porcine preimplantation embryo development by infecting lentivirus carrying exogenous *hOCT4*. Unexpectedly, elevated *XIST* expression but

unchanged expressions of other XIC-linked genes were observed in blastocysts when exogenous *OCT4* was passed to the cleaved embryos. Even though this upregulation of *XIST* might be originated from accelerated lineage segregation in cleaved embryos, the possibility that the functions of porcine *OCT4* on *XIST* expression in preimplantation embryo would be different to that of mouse *Oct4* should be considered. Indeed, one study reported that RNAi-mediated knockdown of *OCT4* didn't influenced *CDX2* expression nevertheless it reduced blastocyst formation rate (Sakurai *et al.* 2013). This might mean that the porcine *OCT4* could regulate its targets differentially in preimplantation embryos compared to that in mouse embryos. As present study couldn't reveal the exact effect of *OCT4* to XIC-linked gene, their relation needs to be confirmed directly in further study using like chromatin-immuoprecipitation (CHIP).

CHAPTER VII

**Correlative comparison for expression levels
between pluripotent genes and XIC-linked genes
in porcine blastocyst**

1. Abstract

X-chromosome inactivation is an epigenetic event with complex molecular regulatory circuits. Pluripotent factors have been suggested to be repressors for XCI because their expressional changes were mutually exclusive with initiation of XCI in developing embryos in mice. However, the insights have been restricted only in mice and the networking in other species is unclear and even any trials were attempted until recently. In previous study, the function of typical pluripotent factor, *OCT4*, on XCI was examined in porcine developing embryos. Subsequently, in here, relations between other represent pluripotent factors, *SOX2*, *REX1*, and *NANOG* and XIC-linked genes were examined by correlative expression analysis in sexed porcine blastocyst. Expression patterns between each gene in two criteria showed similar correlation patterns in male and female blastocysts but expression levels of two pairs of genes, *XIST-NANOG* and *REX1-RLIM*, were correlated oppositely between male and female blastocysts. Although it is hard to conclude that they really related each other, because of weakness of the correlative analysis as an evidence for functions of genes, the result suggest the possibility that various pluripotent factor might be related to XIC-linked genes in pigs.

Key words: Pluripotent gene, X-chromosome inactivation center, Correlation, Preimplantation embryos

2. Introduction

X-chromosome inactivation (XCI) is an essential epigenetic process in eutherians for matching the expression levels of genes in sex chromosome between male and female adults. This process is occurred normally in early embryo generation and known to be induced by non-coding RNA (ncRNA), *XIST*. A number of studies in mice confirmed that the molecular networking for regulating *Xist* expression or its repression was very complicated, and numerous factors were associated with XCI directly or indirectly for concise regulation of the process in developing embryos. Among the various factors, pluripotent factors like *Oct4*, *Sox2*, and *Nanog* were suggested to be typical *trans* regulators which were not located in X-chromosome (Navarro *et al.* 2008). Expression and epigenetic dynamics in developing embryos and embryonic stem (ES) cells, for example, similar time window for *Nanog* expression and erase of imprinted XCI in paternal X in inner cell mass (ICM) of blastocyst (Mak *et al.* 2004), and rapid reduction of typical pluripotent genes and elevation of *Xist* expression in differentiating mouse ES cells (Marahrens *et al.* 1998; Nichols and Smith 2009) were clues that these two epigenetic categories would be linked each other. Indeed, after one study reported by Navarro and his colleagues' in 2008, various pluripotent factors, like *Oct4*, *Sox2*, *Nanog*, *Myc*, *Klf4*, *Rex1*, and *Prdm14* were suggested to regulate XCI by repressing

Xist (*Oct4*, *Sox2*, *Nanog*, *Rex1*, *Prdm14*) (Navarro *et al.* 2008; Donohoe *et al.* 2009; Ma *et al.* 2011; Gontan *et al.* 2012) and its activator, *Rlim* (Navarro *et al.* 2011), or regulating *Tsix* expression (*Oct4*, *Sox2*, *Myc*, *Rex1*, *Klf4*, and *Prdm14*) (Donohoe *et al.* 2009; Navarro *et al.* 2010; Payer *et al.* 2013) in mice. However, the molecular networking between pluripotent factors and XCI has not been examined in any other species except mouse models until recently despite the diversities of the mechanism among the eutherians (Okamoto and Heard 2009; Escamilla-Del-Arenal *et al.* 2011). In previous study, the relation between *OCT4* and X-chromosome inactivation center (XIC)-linked genes was examined by correlative comparison between them in blastocyst, and analyzing effect of *OCT4* overexpression to expression of XIC-linked gene in blastocyst (see the Chapter VI). Even though the exact relation was not confirmed in the study, at least, the opposite relation of the genes with *Xist* was observed compared to that of mice. Indeed, even though the observed expression patterns would be resulted from second hand results, accelerated lineage segregation, the relation between other pluripotent factors and XIC-linked genes needs to be examined to suggest the presence of their association with XCI in pigs.

Therefore, in this study, correlative comparison of expression levels between XIC-linked genes and pluripotent factors, *Sox2*, *Rex1*, and *Nanog*, which were reported to deterring XCI by repressing *Xist* expression rather than activating *Tsix* expression in mice, were conducted in porcine fertilized embryos. Although the analysis only showed weak evidences for the genuine association among them, the

results suggest the possibility that pluripotent factors would be related to XCI also in pigs even though the relation would be opposite to that in mice.

3. Materials and Methods

Ethics statement

The experiments in this study were performed with approval of Institutional Animal Care and Use Committees, Seoul National University (SNU-140325-3).

Production of *in vitro* fertilized blastocyst

Immature oocytes were matured prior to produce the *in vitro* fertilized blastocyst. Detail explanations for the steps were described in Chapter V. Briefly, Cumulus-oocyte-complexes (COCs) were extracted from the ovaries of prepubertal gilts which were donated from Soosan Biotech Research Institute (Seoul, Korea). COCs with multiple cumulus cell layers and granulated cytoplasm were collected and cultured in tissue culture medium (TCM-199; Life Technology, Rockville, MD, USA) containing 10 ng/ml epidermal growth factor (EGF), 1 µg/ml insulin (Sigma-Aldrich, St. Louis, MO, USA) and 10% follicular fluid for 44 hours at 39°C in a 5% CO₂ conditions. Chorionic gonadotropins from humans and equines (hCG, eCG, respectively, 4 IU/ml, Intervet, Boxmeer, Netherlands) were treated for only first 22 hours. Expanded cumulus cells in cultured COCs were denuded using 0.1% of hyaluronidase (Sigma-Aldrich, St. Louis, MO, USA) and followed by fertilized.

Semen from ducor breed (DARBI A.I center, Jochiwon, Korea) was washed with Dulbecco's phosphate buffered saline (DPBS; Welgene, Seoul, Korea) containing 0.1% bovine serum albumin (BSA; Sigma-Aldrich, St.Louis, MO, USA) before use. Prepared sperms concentrated to 1×10^5 cells/ml were treated to the mature oocytes on the modified tris-buffered medium (mTBM) (Abeydeera and Day 1997) and co-incubated for 6 hours at 39°C in a 5% CO₂ conditions. After the incubation, the attached sperms were removed from oocytes by gentle pipetting, followed by being passed to culture medium, porcine zygote medium 3 (PZM3) (Yoshioka *et al.* 2002), and incubated at 39°C in a 5% CO₂ and 5% O₂ conditions.

RNA extraction and Reverse transcription

Messenger RNA was extracted from individual blastocyst by using Dynabeads® mRNA DIRECT™ Kit (Invitrogen, Carlsbad, CA, USA) in accordance with manufacture's guidance. Before mRNA extraction, zona pellucidae of unhatched blastocyst were removed by using Tyrode's acid (Sigma-Aldrich, St. Louis, MO, USA). Extracted mRNA was reverse-transcribed using High Capacity RNA-to-cDNA™ Kit (Applied Biosystems, Foster City, CA, USA) following the manufacturer's guidance.

Quantitative RT-PCR

Quantitative RT-PCR was carried out using DyNAmo HS SYBR Green qPCR kit (Thermo Scientific, Rockford, IL, USA) following the manufacturer's guidance and 0.1 μ M of primer sets (Table VII.1) with 10 μ l reactions. The reaction was carried out following conditions, one cycle of 50°C for 5 minutes; one cycle of 95°C for 5 minutes; 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. *ACTB* and *RN18S* were used for reference gene.

Statistical analysis

Graphpad Prism statistical program (Graphpad Software, San Diego, CA, USA) was used for statistical analysis. Expression levels of target genes in blastocyst following the sex were performed using Student *t*-test. Correlative comparison between expression levels of pluripotent genes and XIC-linked genes was carried out using Pearson correlation coefficient. All data in here were shown as mean \pm standard error mean (S.E.M) and $P < 0.05$ was considered to be significantly different.

Table VII.1. Primer pairs used in quantitative RT-PCR		
Gene symbol	Primer sequence (5 to 3)	Amplicon size (bp)
<i>SOX2</i>	F GCCAGAAGAGGAGGGAAGC	91 bp
	R GCGAGGAAAATCAGACGAAGA	
<i>REX1</i>	F GGACGAAGCCAAGTAAAAATAAAA	120 bp
	R GGGGACTGTGAACGGAGAG	
<i>NANOG</i>	F CTCTCCTCTTCCTTCCTCCA	115 bp
	R TTCCTCCTTGTCTGTGCTCTTC	
<i>XIST</i>	F GCTCCAACCAATCTAAAAGGA	131 bp
	R ATGCCCCATCTCCACCTAA	
<i>LOC102165544</i>	F CTAAGATGGCGGCGTTTG	135 bp
	R TTGTTTTTTCAGGGAATAGAGAGG	
<i>RLIM</i>	F CCCACCACCGCAAAACTC	159 bp
	R CGGCTCACTGCTCTCCAA	
<i>RN18S</i>	F ACAAATCGCTCCACCAACTAAGA	90 bp
	R CGGACACGGACAGGATTGAC	
<i>ACTB</i>	F GTGGACATCAGGAAGGACCTCTA	131 bp
	R ATGATCTTGATCTTCATGGTGCT	

4. Results

Correlative comparison of expression levels between pluripotent genes and XIC-linked genes.

Correlation in expression levels between pluripotent genes, *SOX2*, *NANOG*, and *REX1*, and XIC-linked genes, *XIST*, *LOC102165544*, and *RLIM*, was examined in each sex of individual blastocyst. Before the analysis, sex of individual fertilized blastocyst were determined by *XIST* expression levels as follow previous study (Park *et al.* 2011). Sex-biased expression of three pluripotent genes was checked. Unexpectedly, differential expression levels of *NANOG* were observed but the others didn't show any differences on their expression levels between two groups (Figure VII.1). Correlative comparison of each pair of genes showed generally similar patterns but two gene pairs, *NANOG-XIST* and *REX1-RLIM* showed opposite correlation in male and female blastocysts (Figure VII.2).

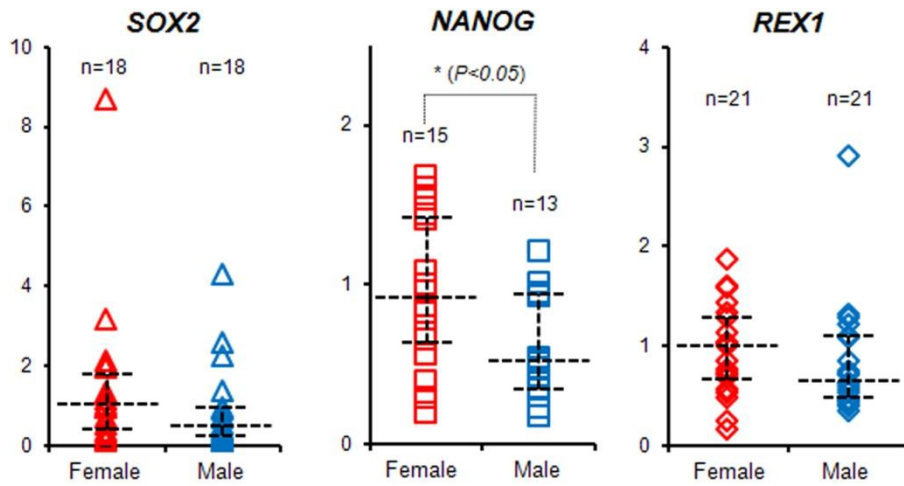


Figure VII.1. Expression levels of pluripotent genes in male and female blastocyst. Expression levels of three pluripotent genes were compared between male and female individual blastocyst sexed by *XIST* expression amount. Each triangle (*SOX2*), square (*NANOG*), and diamond (*REX1*) indicates relative expression levels of genes in individual blastocyst. Median of each gene expression level in female blastocysts was set to one fold. Box plot in each column indicates quartile distribution of expression levels of genes. Asterisk indicates significant differences ($P < 0.05$).

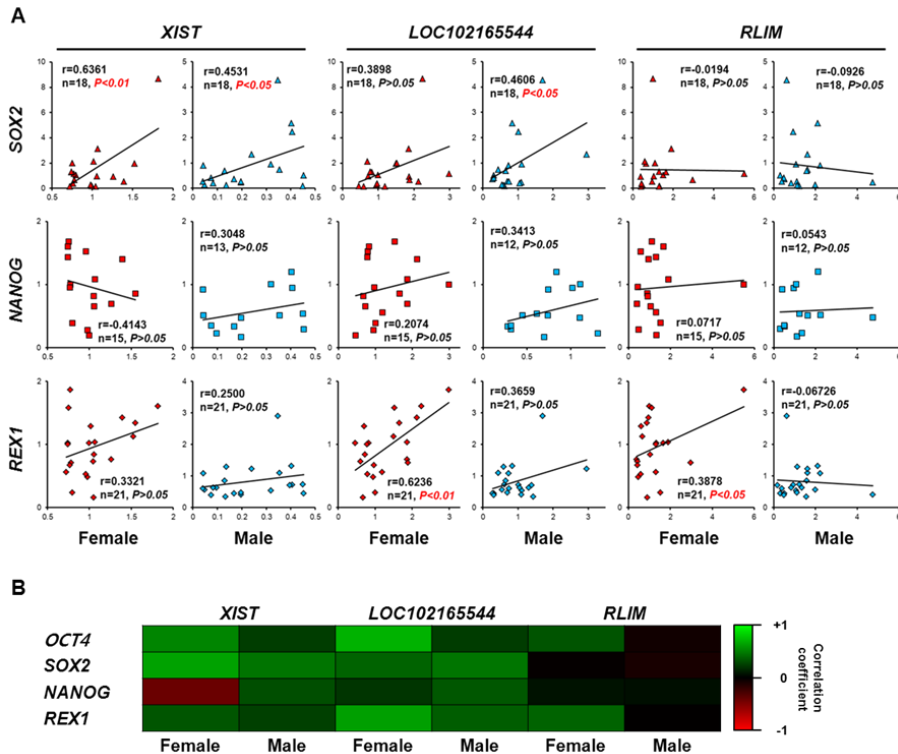


Figure VII.2. Correlative expression analysis between pluripotent genes and XIC-linked genes. (A) Correlative analysis of expression levels of each pluripotent and XIC-linked genes in male and female blastocysts. Each mark on each small panel indicates profiled gene expression levels in individual blastocysts. X- and Y- axis indicate relative expression levels of XIC-linked genes and pluripotent genes, respectively. Median value of each gene expression level in female blastocyst group was considered to one fold. $P < 0.05$ was considered to be significantly correlated. (B) Heatmap representing correlation coefficient of each gene pair. Greenish or reddish colors of boxes indicate positive or negative correlation in expression levels of compared two genes.

5. Discussion

Molecular networking for regulating XCI has been considered to be linked closely to pluripotent markers. Several studies revealed that numerous pluripotent markers would be related to XCI by repressing *XIST* expression. However, the knowledge has been restricted only in mice until recently, and even the examining the genetic relation between expression of *XIST* and pluripotent genes has not been tried in any other species. However, as observed in recent study that the diversities on the XCI process among the species were present (Okamoto *et al.* 2011), the requirement of the analyzing XCI mechanism and its regulators in various species has been highlighted (Escamilla-Del-Arenal *et al.* 2011). As *in vitro* cultured female cell lines carrying two active X-chromosomes like mouse ES cells have not been available in pigs, their parthenogenic blastocysts were used to analyze the genetic relation between *OCT4* and XIC-linked genes alternatively (see the Chapter VI). Successively, correlation in expression levels of other pluripotent genes, *SOX2*, *REX1*, and *NANOG*, with XIC-linked genes, *XIST*, *LOC102165544*, and *RLIM*, in blastocyst was examined in here to suggest the possible association among the genes.

Similar correlative expression patterns were observed between male and female blastocyst except two gene pairs, *NANOG-XIST* and *REX1-RLIM* in this study

(Figure VII.2). Expressions of *NANOG* and *XIST* were negatively and positively correlated in female and male blastocysts, respectively, and those of *REX1* and *RLIM* was oppositely correlated in blastocysts of each sex compared to the relations in *NANOG* and *XIST*. And their correlation coefficient was very mild in male blastocysts. Indeed, correlative analysis of expression level is weak evidences for the genuine relation between two genes, and it is hard to conclude that examined two genes are directly linked to each other. However, considering that XCI is a female specific event and female specifically observed correlation patterns observed in this study (*XIST-NANOG* and *REX1-RLIM*), there might be relation among them. Interestingly, the functional relations of the pairs, negative regulation of *Nanog* to *Xist* and *Rlim* to *Rex1*, were reported in mouse ES cells previously (Navarro *et al.* 2008; Gontan *et al.* 2012). *Nanog* binds to *Xist* intron 1 and its deletion resulted in aberrant *Xist* expression in mouse ES cells, therefore *Nanog* has been considered to repressor for *Xist* expression before XCI (Navarro *et al.* 2008). As previous reports, expression levels comparison in this study showed negatively correlated expression levels between *NANOG* and *XIST* in female blastocyst even though they were not significant. Although it is hard to conclude that porcine *NANOG* really prevent porcine *XIST* expression by binding to its intron as in mouse, the pluripotent factor might regulate *XIST* expression negatively also in pigs. *Rlim*, which ligase E3 ubiquitin to its target proteins and regulate them by degrading, has been reported to target *Rex1* (Gontan *et al.* 2012). *Rex1* is rapidly decreased after XCI initiation in developmental period (Schulz and

Heard 2013) and *Rlim* expression was transiently elevated ES cell differentiation and subsequently reduced by being subjected to XCI (Jonkers *et al.* 2009; Barakat *et al.* 2011). These studies revealed that the timing of expression pattern changes of *Rex1* and *Rlim* overlap each other during mouse XCI. However, the presence of similar relation is unclear in pigs because although correlation was observed strongly in female rather than male blastocyst, the gene expressions in female blastocysts were not negatively correlated (Figure VII.2). And also, one report revealed that *Nanog* repress *Rlim* in mouse ES cells (Navarro *et al.* 2011) but the result didn't showed any relation between *RLIM* and *NANOG*.

As described above, the method used in this study showed only restricted possibilities on the true relation among the compared genes. Nevertheless limits on concluding that the pluripotent factors showing correlative expression with XIC-linked genes in blastocysts are genuine regulator for XCI, some cases showing female specific correlations like *NANOG-XIST* and *REX1-RLIM* in porcine blastocyst, might suggest the possibilities that they could be a factor participate in molecular circuit for XCI in pigs. With the clues of possible relation between pluripotent factors and XIC-linked genes obtained in this study, functional studies on these genes are required further to confirm the relation between pluripotent factors and XCI directly in pigs.

CHAPTER VIII

Conclusion

Conclusion

After the first suggestion of X-chromosome inactivation (XCI) in 1961, huge number of studies has been confirmed the process and molecular mechanisms of the epigenetic phenomenon in mouse embryonic development and embryonic stem (ES) cells. Numerous studies on the process have helped the understanding of the process in developmental biology and as follow increasing the interests on stem cell biology and embryo development, close relation of XCI in the research field has been focused recently. However, although the deep insights of XCI and its regulators has been accumulated for over sixty years, the knowledge on the process is only restricted in mouse models. Indeed, evolutionary diversities on the XCI process and its major regulators, non-coding RNAs (ncRNAs) like *XIST* and *TSIX*, are one of the important features of this epigenetic process, and suggests requirement of species-specific view and studies on the event. Therefore, confirming the process in diverse species is required nevertheless the studies in mouse have been stabilize and set the basis of the knowledge regarding the process efficiently

So in this study, I've identified *XIST* and X-chromosome inactivation center (XIC), which is main genomic region consisting of numerous regulators for XCI, and analyzed the process during embryonic development in pigs of which

importance in biomedical and developmental science has been addressed as a model animal. And the regulation of the process was tried to be confirmed by examining the relation between pluripotent genes, which were suggested to be regulators for XCI and XIC-linked genes.

When I started the studies in pigs, unfortunately, *XIST* ortholog in pigs had not been identified. That's the reason why identification of porcine *XIST* was the first step for this study, as the gene is essential component for XCI. The first study successfully identified reliable porcine *XIST*. Identified porcine *XIST* shared the genomic characters with other orthologs, like female specific expression and its specific promoter methylation on active X-chromosome in fully differentiated cells which were considered generally to be accomplished XCI. And also, the composition of exons and repeat sequences in transcript also similar to its orthologs in other species. Interestingly, the structural and expressional properties were conserved in porcine *XIST* as the orthologs in other eutherians but the sequences of the transcript were extremely variable among them. This structural but not sequence conservation was also observed in several repeat regions in porcine *XIST*. Some repeat regions in porcine *XIST*, which were considered to be originated from mobile elements, shared similar length of their monomers with those in non-rodent eutherians, but the their sequences were not conserved among them. This was expected to be the common property of ncRNAs in XIC which were evolved

rapidly in eutherians. This structural similarity with less conserved sequences among the orthologs was also observed in other ncRNAs in XIC.

In second study, porcine XIC was identified by comparative genomic analysis and it was considered to be syntenic with that in other species. Although the uncharacterized genes in porcine XIC showed low sequence similarity with their expected counterpart ncRNAs in human and mouse XIC, they would be orthologs with the counterparts when considering the evolutionary conserved features, like composition of the genomic region flanking ncRNAs loci and transcription strand of each gene. Consequently it is thought that pigs have conserved XIC but each compartment coded on the region, especially ncRNAs, are less sequence conserved among the eutherians. The less conserved sequences of the ncRNAs in porcine XIC would mean that the XIC-borne ncRNAs might have differential functions in XCI, or other physiological process would be present in pigs compared to those in mice. Expression patterns of XIC-linked genes in porcine developing embryos showed the possibility that pigs have their own time window and mechanism for XCI differing to those of mice. Dramatic increase of *XIST* was observed after morula stage and rapid reduction of two XIC-linked protein coding genes, *RLIM* and *CHIC1*, would mean dosage compensation of the gene is achieved already in blastocysts. This also supported by compensated expression of two genes in male and female blastocyst and single allelic methylation of *CHIC1* promoter in female

blastocyst. However, demethylation of *XIST* promoter was observed in female blastocyst and this is different with that in female somatic cells. Considering the stabilization of XCI required about one week after initiation of *Xist* coating in differentiating mouse ES cells, and presence of the intermediate XCI status which shows silenced genes and exchanged histone marks to heterochromatin without coated methylation on DNA, the results in this study would be concluded that dosage compensation of several X-linked genes would be initiated in or before blastocyst, and stabilized XCI would be achieved developmentally later stages in pigs. And also, non-imprinted biallelic expression of *XIST* raised the possibility that XCI process in pigs would differ to that of mouse preimplantation embryos showing two step inactivation, imprinted and random XCI. Therefore, it is considered that pigs would have their specific process for XCI and its initiation timing.

Although it was not clear, pluripotent factors might be related to XCI in pigs. In the fourth and fifth study, correlative expressions between represent pluripotent factors and XCI-linked genes, which were suggested to be associated with XCI mouse model, were observed in porcine female blastocysts. Unexpectedly, positively correlated expression patterns, except *NANOG* and *XIST*, were observed in female blastocysts and these results might mean to non-repressive actions of pluripotent factors to XCI. Because the pluripotent factors are associated with and

regulate numerous downstream targets, it was hard to determine clear effect of pluripotent factors to XCI in this study. Overexpression of *OCT4* in preimplantation embryos resulted in *XIST* elevation. Exogenous *OCT4* would induced disruption of molecular networking and this might resulted in accelerating differentiation, follow by *XIST* increase in porcine embryos. Interestingly, this unexpected *XIST* elevation in female blastocysts was shown when the exogenous gene was passed to cleaved embryos rather than early blastocysts. Considering *XIST* expression was explosively increased after morulae and exclusively mutual expression of pluripotent genes and *Xist* in mouse epiblast, it is possible that segregation of lineages might be initiated considerably early stage in pigs. However, all of them are possibilities, therefore, additional studies confirming the roles of pluripotent factors in XCI is still required for the comprehensive understanding of the complex molecular networking in pigs.

In here, I've identified and annotated XCI-related genes and observed changes on dosages of X-linked genes in porcine developing embryos. Although the numerous studies would be required additionally to clearly know this evolutionary diverse epigenetic event in pigs, in conclusion, pigs would have their own process for the XCI which is distinct to mouse species. This study would be basics for the XCI studies in pigs.

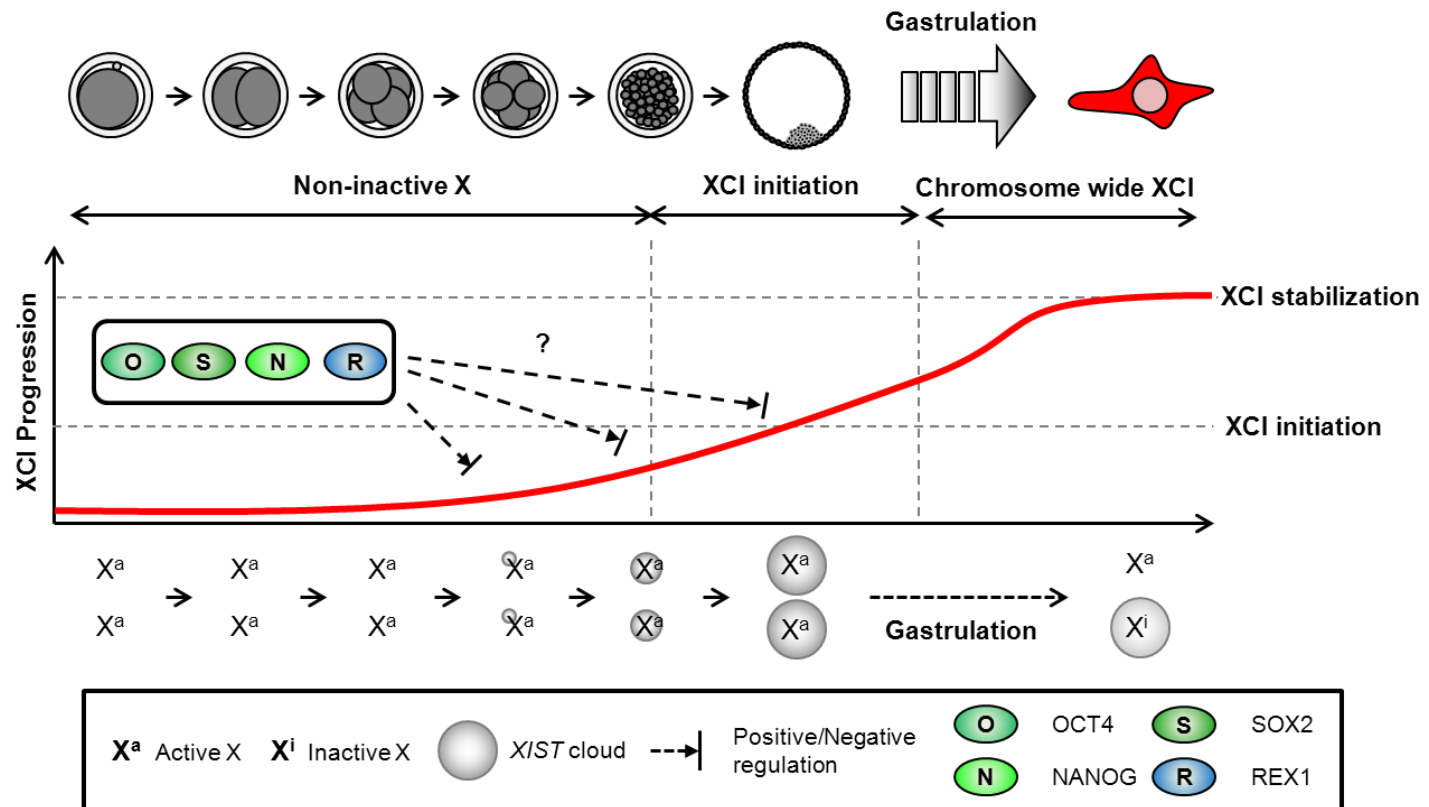


Figure VIII. XCI in porcine embryo development.

References

Abeydeera, L.R., and Day, B.N. (1997) In vitro penetration of pig oocytes in a modified Tris-buffered medium: effect of BSA, caffeine and calcium. *Theriogenology* 48(4), 537-44

Asadi, M.H., Mowla, S.J., Fathi, F., Aleyasin, A., Asadzadeh, J., and Atlasi, Y. (2011) OCT4B1, a novel spliced variant of OCT4, is highly expressed in gastric cancer and acts as an antiapoptotic factor. *Int J Cancer* 128(11), 2645-52

Atlasi, Y., Mowla, S.J., Ziaee, S.A., Gokhale, P.J., and Andrews, P.W. (2008) OCT4 spliced variants are differentially expressed in human pluripotent and nonpluripotent cells. *Stem Cells* 26(12), 3068-74

Augui, S., Nora, E.P., and Heard, E. (2011) Regulation of X-chromosome inactivation by the X-inactivation centre. *Nat Rev Genet* 12(6), 429-42

Barakat, T.S., Gunhanlar, N., Pardo, C.G., Achame, E.M., Ghazvini, M., Boers, R., Kenter, A., Rentmeester, E., Grootegoed, J.A., and Gribnau, J. (2011) RNF12 activates Xist and is essential for X chromosome inactivation. *PLoS Genet* 7(1), e1002001

Barakat, T.S., Loos, F., van Staveren, S., Myronova, E., Ghazvini, M., Grootegoed, J.A., and Gribnau, J. (2014) The trans-activator RNF12 and cis-acting elements effectuate X chromosome inactivation independent of X-pairing. *Mol Cell* 53(6), 965-78

Baumann, C., and De La Fuente, R. (2009) ATRX marks the inactive X chromosome (Xi) in somatic cells and during imprinted X chromosome inactivation in trophoblast stem cells. *Chromosoma* 118(2), 209-22

Bermejo-Alvarez, P., Rizos, D., Lonergan, P., and Gutierrez-Adan, A. (2011) Transcriptional sexual dimorphism in elongating bovine embryos: implications for XCI and sex determination genes. *Reproduction* 141(6), 801-8

Bermejo-Alvarez, P., Rizos, D., Rath, D., Lonergan, P., and Gutierrez-Adan, A. (2010) Sex determines the expression level of one third of the actively expressed genes in bovine blastocysts. *Proc Natl Acad Sci U S A* 107(8), 3394-9

Bischoff, S.R., Tsai, S.Q., Hardison, N.E., Motsinger-Reif, A.A., Freking, B.A., Nonneman, D.J., Rohrer, G.A., and Piedrahita, J.A. (2013) Differences in X-Chromosome Transcriptional Activity and Cholesterol Metabolism between Placentae from Swine Breeds from Asian and Western Origins. *PLoS ONE* 8(1), e55345

Blewitt, M.E., Gendrel, A.V., Pang, Z., Sparrow, D.B., Whitelaw, N., Craig, J.M., Apedaile, A., Hilton, D.J., Dunwoodie, S.L., Brockdorff, N., Kay, G.F., and Whitelaw, E. (2008) SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat Genet* 40(5), 663-9

Bouniol, C., Nguyen, E., and Debey, P. (1995) Endogenous transcription occurs at the 1-cell stage in the mouse embryo. *Exp Cell Res* 218(1), 57-62

Braude, P., Bolton, V., and Moore, S. (1988) Human gene expression first occurs between the four- and eight-cell stages of preimplantation development. *Nature* 332(6163), 459-61

Brockdorff, N., Ashworth, A., Kay, G.F., Cooper, P., Smith, S., McCabe, V.M., Norris, D.P., Penny, G.D., Patel, D., and Rastan, S. (1991) Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* 351(6324), 329-31

Brockdorff, N., Ashworth, A., Kay, G.F., McCabe, V.M., Norris, D.P., Cooper, P.J., Swift, S., and Rastan, S. (1992) The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* 71(3), 515-26

Brons, I.G., Smithers, L.E., Trotter, M.W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., and Vallier, L. (2007) Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 448(7150), 191-5

Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R., and Willard, H.F. (1991a) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* 349(6304), 38-44

Brown, C.J., Hendrich, B.D., Rupert, J.L., Lafreniere, R.G., Xing, Y., Lawrence, J., and Willard, H.F. (1992) The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* 71(3), 527-42

Brown, C.J., Lafreniere, R.G., Powers, V.E., Sebastio, G., Ballabio, A., Pettigrew, A.L., Ledbetter, D.H., Levy, E., Craig, I.W., and Willard, H.F. (1991b) Localization of the X inactivation centre on the human X chromosome in Xq13. *Nature* 349(6304), 82-4

Brown, C.J., and Robinson, W.P. (1997) XIST expression and X-chromosome inactivation in human preimplantation embryos. *Am J Hum Genet* 61(1), 5-8

Brown, S.D. (1991) XIST and the mapping of the X chromosome inactivation centre. *Bioessays* 13(11), 607-12

Cauffman, G., Liebaers, I., Van Steirteghem, A., and Van de Velde, H. (2006) POU5F1 isoforms show different expression patterns in human embryonic stem cells and preimplantation embryos. *Stem Cells* 24(12), 2685-91

Cepica, S., Masopust, M., Knoll, A., Bartenschlager, H., Yerle, M., Rohrer, G.A., and Geldermann, H. (2006) Linkage and RH mapping of 10 genes to a QTL region for fatness and muscling traits on pig chromosome X. *Anim Genet* 37(6), 603-4

Chadwick, B.P., Valley, C.M., and Willard, H.F. (2001) Histone variant macroH2A contains two distinct macrochromatin domains capable of directing macroH2A to the inactive X chromosome. *Nucleic Acids Res* 29(13), 2699-705

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003) Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* 113(5), 643-55

Chang, S.C., and Brown, C.J. (2010) Identification of regulatory elements flanking human XIST reveals species differences. *BMC Mol Biol* 11, 20

Chao, W., Huynh, K.D., Spencer, R.J., Davidow, L.S., and Lee, J.T. (2002) CTCF, a candidate trans-acting factor for X-inactivation choice. *Science* 295(5553), 345-7

Chaumeil, J., Le Baccon, P., Wutz, A., and Heard, E. (2006) A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev* 20(16), 2223-37

Chaumeil, J., Okamoto, I., Guggiari, M., and Heard, E. (2002) Integrated kinetics of X

chromosome inactivation in differentiating embryonic stem cells. *Cytogenet Genome Res* 99(1-4), 75-84

Chow, J., and Heard, E. (2009) X inactivation and the complexities of silencing a sex chromosome. *Curr Opin Cell Biol* 21(3), 359-66

Chureau, C., Chantalat, S., Romito, A., Galvani, A., Duret, L., Avner, P., and Rougeulle, C. (2011) Ftx is a non-coding RNA which affects Xist expression and chromatin structure within the X-inactivation center region. *Hum Mol Genet* 20(4), 705-18

Chureau, C., Prissette, M., Bourdet, A., Barbe, V., Cattolico, L., Jones, L., Eggen, A., Avner, P., and Duret, L. (2002) Comparative sequence analysis of the X-inactivation center region in mouse, human, and bovine. *Genome Res* 12(6), 894-908

Clerc, P., and Avner, P. (1998) Role of the region 3' to Xist exon 6 in the counting process of X-chromosome inactivation. *Nat Genet* 19(3), 249-53

Costanzi, C., and Pehrson, J.R. (1998) Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* 393(6685), 599-601

Daniels, R., Zuccotti, M., Kinis, T., Serhal, P., and Monk, M. (1997) XIST expression in human oocytes and preimplantation embryos. *Am J Hum Genet* 61(1), 33-9

De La Fuente, R., Hahnel, A., Basrur, P.K., and King, W.A. (1999) X inactive-specific transcript (Xist) expression and X chromosome inactivation in the preattachment bovine embryo. *Biol Reprod* 60(3), 769-75

Deakin, J.E., Chaumeil, J., Hore, T.A., and Marshall Graves, J.A. (2009) Unravelling the evolutionary origins of X chromosome inactivation in mammals: insights from marsupials and monotremes. *Chromosome Res* 17(5), 671-85

Deakin, J.E., Hore, T.A., Koina, E., and Marshall Graves, J.A. (2008) The status of dosage compensation in the multiple X chromosomes of the platypus. *PLoS Genet* 4(7), e1000140

Debrand, E., Chureau, C., Arnaud, D., Avner, P., and Heard, E. (1999) Functional analysis of the DXPas34 locus, a 3' regulator of Xist expression. *Mol Cell Biol* 19(12), 8513-25

Dindot, S.V., Farin, P.W., Farin, C.E., Romano, J., Walker, S., Long, C., and Piedrahita, J.A. (2004a) Epigenetic and genomic imprinting analysis in nuclear transfer derived Bos gaurus/Bos taurus hybrid fetuses. *Biol Reprod* 71(2), 470-8

Dindot, S.V., Kent, K.C., Evers, B., Loskutoff, N., Womack, J., and Piedrahita, J.A. (2004b) Conservation of genomic imprinting at the XIST, IGF2, and GTL2 loci in the bovine. *Mamm Genome* 15(12), 966-74

Donohoe, M.E., Silva, S.S., Pinter, S.F., Xu, N., and Lee, J.T. (2009) The pluripotency factor Oct4 interacts with Ctfc and also controls X-chromosome pairing and counting. *Nature* 460(7251), 128-32

Donohoe, M.E., Zhang, L.F., Xu, N., Shi, Y., and Lee, J.T. (2007) Identification of a Ctfc cofactor, Yy1, for the X chromosome binary switch. *Mol Cell* 25(1), 43-56

Doyen, C.M., An, W., Angelov, D., Bondarenko, V., Miettton, F., Studitsky, V.M.,

Hamiche, A., Roeder, R.G., Bouvet, P., and Dimitrov, S. (2006) Mechanism of polymerase II transcription repression by the histone variant macroH2A. *Mol Cell Biol* 26(3), 1156-64

Dupont, C., and Gribnau, J. (2013) Different flavors of X-chromosome inactivation in mammals. *Curr Opin Cell Biol* 25(3), 314-21

Duret, L., Chureau, C., Samain, S., Weissenbach, J., and Avner, P. (2006) The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. *Science* 312(5780), 1653-5

Eggan, K., Akutsu, H., Hochedlinger, K., Rideout, W., 3rd, Yanagimachi, R., and Jaenisch, R. (2000) X-Chromosome inactivation in cloned mouse embryos. *Science* 290(5496), 1578-81

Elisaphenko, E.A., Kolesnikov, N.N., Shevchenko, A.I., Rogozin, I.B., Nesterova, T.B., Brockdorff, N., and Zakian, S.M. (2008) A dual origin of the Xist gene from a protein-coding gene and a set of transposable elements. *PLoS One* 3(6), e2521

Escamilla-Del-Arenal, M., da Rocha, S.T., and Heard, E. (2011) Evolutionary diversity and developmental regulation of X-chromosome inactivation. *Hum Genet* 130(2), 307-27

Ewerling, S., Hofmann, A., Klose, R., Weppert, M., Brem, G., Rink, K., Pfeifer, A., and Wolf, E. (2006) Evaluation of laser-assisted lentiviral transgenesis in bovine. *Transgenic Res* 15(4), 447-54

Eymery, A., Callanan, M., and Vourc'h, C. (2009) The secret message of heterochromatin: new insights into the mechanisms and function of centromeric and pericentric repeat

sequence transcription. *Int J Dev Biol* 53(2-3), 259-68

Farashahi Yazd, E., Rafiee, M.R., Soleimani, M., Tavallaei, M., Salmani, M.K., and Mowla, S.J. (2011) OCT4B1, a novel spliced variant of OCT4, generates a stable truncated protein with a potential role in stress response. *Cancer Lett* 309(2), 170-5

Frei, R.E., Schultz, G.A., and Church, R.B. (1989) Qualitative and quantitative changes in protein synthesis occur at the 8-16-cell stage of embryogenesis in the cow. *J Reprod Fertil* 86(2), 637-41

Fujishiro, S.H., Nakano, K., Mizukami, Y., Azami, T., Arai, Y., Matsunari, H., Ishino, R., Nishimura, T., Watanabe, M., Abe, T., Furukawa, Y., Umeyama, K., Yamanaka, S., Ema, M., Nagashima, H., and Hanazono, Y. (2013) Generation of naive-like porcine-induced pluripotent stem cells capable of contributing to embryonic and fetal development. *Stem Cells Dev* 22(3), 473-82

Funahashi, H., Cantley, T.C., and Day, B.N. (1997) Synchronization of meiosis in porcine oocytes by exposure to dibutyl cyclic adenosine monophosphate improves developmental competence following in vitro fertilization. *Biol Reprod* 57(1), 49-53

Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A., Rais, Y., Shipony, Z., Mukamel, Z., Krupalnik, V., Zerbib, M., Geula, S., Caspi, I., Schneir, D., Schwartz, T., Gilad, S., Amann-Zalcenstein, D., Benjamin, S., Amit, I., Tanay, A., Massarwa, R., Novershtern, N., and Hanna, J.H. (2013) Derivation of novel human ground state naive pluripotent stem cells. *Nature* 504(7479), 282-6

Gao, Y., Hyttel, P., and Hall, V.J. (2011) Dynamic changes in epigenetic marks and gene expression during porcine epiblast specification. *Cell Reprogram* 13(4), 345-60

Gillich, A., Bao, S., Grabole, N., Hayashi, K., Trotter, M.W., Pasque, V., Magnusdottir, E., and Surani, M.A. (2012) Epiblast stem cell-based system reveals reprogramming synergy of germline factors. *Cell Stem Cell* 10(4), 425-39

Gontan, C., Achame, E.M., Demmers, J., Barakat, T.S., Rentmeester, E., van, I.W., Grootegoed, J.A., and Gribnau, J. (2012) RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. *Nature* 485(7398), 386-90

Goto, T., Wright, E., and Monk, M. (1997) Paternal X-chromosome inactivation in human trophoblastic cells. *Mol Hum Reprod* 3(1), 77-80

Grant, J., Mahadevaiah, S.K., Khil, P., Sangrithi, M.N., Royo, H., Duckworth, J., McCarrey, J.R., VandeBerg, J.L., Renfree, M.B., Taylor, W., Elgar, G., Camerini-Otero, R.D., Gilchrist, M.J., and Turner, J.M. (2012) Rxs is a metatherian RNA with Xist-like properties in X-chromosome inactivation. *Nature* 487(7406), 254-8

Groenen, M.A., Archibald, A.L., Uenishi, H., Tuggle, C.K., Takeuchi, Y., Rothschild, M.F., Rogel-Gaillard, C., Park, C., Milan, D., Megens, H.J., Li, S., Larkin, D.M., Kim, H., Frantz, L.A., Caccamo, M., Ahn, H., Aken, B.L., Anselmo, A., Anthon, C., Auvil, L., Badaoui, B., Beattie, C.W., Bendixen, C., Berman, D., Blecha, F., Blomberg, J., Bolund, L., Bosse, M., Botti, S., Bujie, Z., Bystrom, M., Capitanu, B., Carvalho-Silva, D., Chardon, P., Chen, C., Cheng, R., Choi, S.H., Chow, W., Clark, R.C., Clee, C., Crooijmans, R.P., Dawson, H.D., Dehais, P., De Sapio, F., Dibbits, B., Drou, N., Du, Z.Q., Eversole, K., Fadista, J., Fairley, S., Faraut, T., Faulkner, G.J., Fowler, K.E., Fredholm, M., Fritz, E., Gilbert, J.G., Giuffra, E., Gorodkin, J., Griffin, D.K., Harrow, J.L., Hayward, A., Howe, K., Hu, Z.L., Humphray, S.J., Hunt, T., Hornshoj, H., Jeon, J.T., Jern, P., Jones, M., Jurka, J., Kanamori, H., Kapetanovic, R., Kim, J., Kim, J.H., Kim, K.W., Kim, T.H., Larson, G., Lee, K., Lee, K.T., Leggett, R., Lewin, H.A., Li, Y., Liu, W., Loveland, J.E., Lu, Y., Lunney, J.K., Ma, J., Madsen, O., Mann, K., Matthews,

L., McLaren, S., Morozumi, T., Murtaugh, M.P., Narayan, J., Nguyen, D.T., Ni, P., Oh, S.J., Onteru, S., Panitz, F., Park, E.W., Park, H.S., Pascal, G., Paudel, Y., Perez-Enciso, M., Ramirez-Gonzalez, R., Reecy, J.M., Rodriguez-Zas, S., Rohrer, G.A., Rund, L., Sang, Y., Schachtschneider, K., Schraiber, J.G., Schwartz, J., Scobie, L., Scott, C., Searle, S., Servin, B., Southey, B.R., Sperber, G., Stadler, P., Sweedler, J.V., Tafer, H., Thomsen, B., Wali, R., Wang, J., White, S., Xu, X., Yerle, M., Zhang, G., Zhang, J., Zhao, S., Rogers, J., Churcher, C., and Schook, L.B. (2012) Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* 491(7424), 393-8

Grutzner, F., Rens, W., Tsend-Ayush, E., El-Mogharbel, N., O'Brien, P.C., Jones, R.C., Ferguson-Smith, M.A., and Marshall Graves, J.A. (2004) In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. *Nature* 432(7019), 913-7

Guo, C.L., Liu, L., Jia, Y.D., Zhao, X.Y., Zhou, Q., and Wang, L. (2012) A novel variant of Oct3/4 gene in mouse embryonic stem cells. *Stem Cell Res* 9(2), 69-76

Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., and Smith, A. (2009) Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* 136(7), 1063-9

Gutierrez-Adan, A., Oter, M., Martinez-Madrid, B., Pintado, B., and De La Fuente, J. (2000) Differential expression of two genes located on the X chromosome between male and female in vitro-produced bovine embryos at the blastocyst stage. *Mol Reprod Dev* 55(2), 146-51

Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat, J., Carey, B.W., and Jaenisch, R. (2010) Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proc Natl Acad*

Sci U S A 107(20), 9222-7

Harrison, K.B. (1989) X-chromosome inactivation in the human cytotrophoblast. *Cytogenet Cell Genet* 52(1-2), 37-41

Hasegawa, Y., Brockdorff, N., Kawano, S., Tsutui, K., and Nakagawa, S. (2010) The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. *Dev Cell* 19(3), 469-76

Heard, E. (2004) Recent advances in X-chromosome inactivation. *Curr Opin Cell Biol* 16(3), 247-55

Heard, E., Rougeulle, C., Arnaud, D., Avner, P., Allis, C.D., and Spector, D.L. (2001) Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. *Cell* 107(6), 727-38

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., Wang, W., Weng, Z., Green, R.D., Crawford, G.E., and Ren, B. (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* 39(3), 311-8

Hendrich, B.D., Brown, C.J., and Willard, H.F. (1993) Evolutionary conservation of possible functional domains of the human and murine XIST genes. *Hum Mol Genet* 2(6), 663-72

Hendrich, B.D., Plenge, R.M., and Willard, H.F. (1997) Identification and characterization of the human XIST gene promoter: implications for models of X chromosome inactivation. *Nucleic Acids Res* 25(13), 2661-71

Hore, T.A., Koina, E., Wakefield, M.J., and Marshall Graves, J.A. (2007) The region homologous to the X-chromosome inactivation centre has been disrupted in marsupial and monotreme mammals. *Chromosome Res* 15(2), 147-61

Hornecker, J.L., Samollow, P.B., Robinson, E.S., Vandeberg, J.L., and McCarrey, J.R. (2007) Meiotic sex chromosome inactivation in the marsupial *Monodelphis domestica*. *Genesis* 45(11), 696-708

Huynh, K.D., and Lee, J.T. (2003) Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. *Nature* 426(6968), 857-62

Hwang, J.Y., Kim, E.B., Ka, H., and Lee, C.K. (2013a) Identification of the porcine XIST gene and its differential CpG methylation status in male and female pig cells. *PLoS One* 8(9), e73677

Hwang, J.Y., Mulligan, B.P., Kim, H.M., Yang, B.C., and Lee, C.K. (2013b) Quantitative analysis of sperm mRNA in the pig: relationship with early embryo development and capacitation. *Reprod Fertil Dev* 25(5), 807-17

Hwang, J.Y., Oh, J.N., Lee, D.K., Choi, K.H., Park, C.H., and Lee, C.K. (2014) Identification and differential expression patterns of porcine OCT4 variants. *Reproduction*

Inoue, K., Kohda, T., Sugimoto, M., Sado, T., Ogonuki, N., Matoba, S., Shiura, H., Ikeda, R., Mochida, K., Fujii, T., Sawai, K., Otte, A.P., Tian, X.C., Yang, X., Ishino, F., Abe, K., and Ogura, A. (2010) Impeding Xist expression from the active X chromosome improves mouse somatic cell nuclear transfer. *Science* 330(6003), 496-9

Jarrell, V.L., Day, B.N., and Prather, R.S. (1991) The transition from maternal to zygotic control of development occurs during the 4-cell stage in the domestic pig, *Sus scrofa*: quantitative and qualitative aspects of protein synthesis. *Biol Reprod* 44(1), 62-8

Jeon, Y., Sarma, K., and Lee, J.T. (2012) New and Xisting regulatory mechanisms of X chromosome inactivation. *Curr Opin Genet Dev* 22(2), 62-71

Jiang, L., Lai, L., Samuel, M., Prather, R.S., Yang, X., and Tian, X.C. (2008) Expression of X-linked genes in deceased neonates and surviving cloned female piglets. *Mol Reprod Dev* 75(2), 265-73

Johnston, C.M., Nesterova, T.B., Formstone, E.J., Newall, A.E., Duthie, S.M., Sheardown, S.A., and Brockdorff, N. (1998) Developmentally regulated Xist promoter switch mediates initiation of X inactivation. *Cell* 94(6), 809-17

Johnston, C.M., Newall, A.E., Brockdorff, N., and Nesterova, T.B. (2002) Enox, a novel gene that maps 10 kb upstream of Xist and partially escapes X inactivation. *Genomics* 80(2), 236-44

Jonkers, I., Barakat, T.S., Achame, E.M., Monkhorst, K., Kenter, A., Rentmeester, E., Grosveld, F., Grootegoed, J.A., and Gribnau, J. (2009) RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation. *Cell* 139(5), 999-1011

Kalantry, S., Purushothaman, S., Bowen, R.B., Starmer, J., and Magnuson, T. (2009) Evidence of Xist RNA-independent initiation of mouse imprinted X-chromosome inactivation. *Nature* 460(7255), 647-51

Kaneko, S., Li, G., Son, J., Xu, C.F., Margueron, R., Neubert, T.A., and Reinberg, D.

(2010) Phosphorylation of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to ncRNA. *Genes Dev* 24(23), 2615-20

Kanhere, A., Viiri, K., Araujo, C.C., Rasaiyaah, J., Bouwman, R.D., Whyte, W.A., Pereira, C.F., Brookes, E., Walker, K., Bell, G.W., Pombo, A., Fisher, A.G., Young, R.A., and Jenner, R.G. (2010) Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2. *Mol Cell* 38(5), 675-88

Kaslow, D.C., and Migeon, B.R. (1987) DNA methylation stabilizes X chromosome inactivation in eutherians but not in marsupials: evidence for multistep maintenance of mammalian X dosage compensation. *Proc Natl Acad Sci U S A* 84(17), 6210-4

Kay, G.F., Barton, S.C., Surani, M.A., and Rastan, S. (1994) Imprinting and X chromosome counting mechanisms determine Xist expression in early mouse development. *Cell* 77(5), 639-50

Kay, G.F., Penny, G.D., Patel, D., Ashworth, A., Brockdorff, N., and Rastan, S. (1993) Expression of Xist during mouse development suggests a role in the initiation of X chromosome inactivation. *Cell* 72(2), 171-82

Koh, S., and Piedrahita, J.A. (2014) From "ES-like" cells to induced pluripotent stem cells: a historical perspective in domestic animals. *Theriogenology* 81(1), 103-11

Kurimoto, K., Yamaji, M., Seki, Y., and Saitou, M. (2008) Specification of the germ cell lineage in mice: a process orchestrated by the PR-domain proteins, Blimp1 and Prdm14. *Cell Cycle* 7(22), 3514-8

Langmead, B., and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat*

Methods 9(4), 357-9

Lee, E., Lee, S.H., Kim, S., Jeong, Y.W., Kim, J.H., Koo, O.J., Park, S.M., Hashem, M.A., Hossein, M.S., Son, H.Y., Lee, C.K., Hwang, W.S., Kang, S.K., and Lee, B.C. (2006a) Analysis of nuclear reprogramming in cloned miniature pig embryos by expression of Oct-4 and Oct-4 related genes. *Biochem Biophys Res Commun* 348(4), 1419-28

Lee, J., Kim, H.K., Rho, J.Y., Han, Y.M., and Kim, J. (2006b) The human OCT-4 isoforms differ in their ability to confer self-renewal. *J Biol Chem* 281(44), 33554-65

Lee, J.T. (2005) Regulation of X-chromosome counting by Tsix and Xite sequences. *Science* 309(5735), 768-71

Lee, J.T., Davidow, L.S., and Warshawsky, D. (1999) Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat Genet* 21(4), 400-4

Lee, J.T., and Lu, N. (1999) Targeted mutagenesis of Tsix leads to nonrandom X inactivation. *Cell* 99(1), 47-57

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16), 2078-9

Liedtke, S., Enczmann, J., Wacławczyk, S., Wernet, P., and Kogler, G. (2007) Oct4 and its pseudogenes confuse stem cell research. *Cell Stem Cell* 1(4), 364-6

Lock, L.F., Takagi, N., and Martin, G.R. (1987) Methylation of the Hprt gene on the

inactive X occurs after chromosome inactivation. *Cell* 48(1), 39-46

Lyon, M.F. (1961) Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* 190, 372-3

Lyon, M.F. (1962) Sex chromatin and gene action in the mammalian X-chromosome. *Am J Hum Genet* 14, 135-48

Ma, Z., Swigut, T., Valouev, A., Rada-Iglesias, A., and Wysocka, J. (2011) Sequence-specific regulator Prdm14 safeguards mouse ESCs from entering extraembryonic endoderm fates. *Nat Struct Mol Biol* 18(2), 120-7

Magnani, L., and Cabot, R.A. (2008) In vitro and in vivo derived porcine embryos possess similar, but not identical, patterns of Oct4, Nanog, and Sox2 mRNA expression during cleavage development. *Mol Reprod Dev* 75(12), 1726-35

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R., Plath, K., and Hochedlinger, K. (2007) Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* 1(1), 55-70

Mak, W., Nesterova, T.B., de Napoles, M., Appanah, R., Yamanaka, S., Otte, A.P., and Brockdorff, N. (2004) Reactivation of the paternal X chromosome in early mouse embryos. *Science* 303(5658), 666-9

Manes, C. (1973) The participation of the embryonic genome during early cleavage in the rabbit. *Dev Biol* 32(2), 453-9

Marahrens, Y., Loring, J., and Jaenisch, R. (1998) Role of the Xist gene in X chromosome choosing. *Cell* 92(5), 657-64

Marahrens, Y., Panning, B., Dausman, J., Strauss, W., and Jaenisch, R. (1997) Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev* 11(2), 156-66

Margueron, R., and Reinberg, D. (2011) The Polycomb complex PRC2 and its mark in life. *Nature* 469(7330), 343-9

Marks, H., Chow, J.C., Denissov, S., Francoijs, K.J., Brockdorff, N., Heard, E., and Stunnenberg, H.G. (2009) High-resolution analysis of epigenetic changes associated with X inactivation. *Genome Res* 19(8), 1361-73

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A.A., Ko, M.S., and Niwa, H. (2007) Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol* 9(6), 625-35

Matoba, S., Inoue, K., Kohda, T., Sugimoto, M., Mizutani, E., Ogonuki, N., Nakamura, T., Abe, K., Nakano, T., Ishino, F., and Ogura, A. (2011) RNAi-mediated knockdown of Xist can rescue the impaired postimplantation development of cloned mouse embryos. *Proc Natl Acad Sci U S A* 108(51), 20621-6

McDonald, L.E., Paterson, C.A., and Kay, G.F. (1998) Bisulfite genomic sequencing-derived methylation profile of the xist gene throughout early mouse development. *Genomics* 54(3), 379-86

McMahon, A., and Monk, M. (1983) X-chromosome activity in female mouse embryos heterozygous for Pgk-1 and Searle's translocation, T(X; 16) 16H. *Genet Res* 41(1), 69-83

Meyer, B.J. (1999) Sex and death of a worm: assessing and repressing X chromosomes. *Harvey Lect* 95, 85-105

Mietton, F., Sengupta, A.K., Molla, A., Picchi, G., Barral, S., Heliot, L., Grange, T., Wutz, A., and Dimitrov, S. (2009) Weak but uniform enrichment of the histone variant macroH2A1 along the inactive X chromosome. *Mol Cell Biol* 29(1), 150-6

Migeon, B.R., Chowdhury, A.K., Dunston, J.A., and McIntosh, I. (2001) Identification of TSIX, encoding an RNA antisense to human XIST, reveals differences from its murine counterpart: implications for X inactivation. *Am J Hum Genet* 69(5), 951-60

Migeon, B.R., Lee, C.H., Chowdhury, A.K., and Carpenter, H. (2002) Species differences in TSIX/Tsix reveal the roles of these genes in X-chromosome inactivation. *Am J Hum Genet* 71(2), 286-93

Milne, I., Bayer, M., Cardle, L., Shaw, P., Stephen, G., Wright, F., and Marshall, D. (2010) Tablet--next generation sequence assembly visualization. *Bioinformatics* 26(3), 401-2

Minkovsky, A., Barakat, T.S., Sellami, N., Chin, M.H., Gunhanlar, N., Gribnau, J., and Plath, K. (2013) The pluripotency factor-bound intron 1 of Xist is dispensable for X chromosome inactivation and reactivation in vitro and in vivo. *Cell Rep* 3(3), 905-18

Mirzaei, M.R., Najafi, A., Arababadi, M.K., Asadi, M.H., and Mowla, S.J. (2014)

Altered expression of apoptotic genes in response to OCT4B1 suppression in human tumor cell lines. *Tumour Biol*

Mizuno, N., and Kosaka, M. (2008) Novel variants of Oct-3/4 gene expressed in mouse somatic cells. *J Biol Chem* 283(45), 30997-1004

Monk, M., and Harper, M.I. (1979) Sequential X chromosome inactivation coupled with cellular differentiation in early mouse embryos. *Nature* 281(5729), 311-3

Moreira de Mello, J.C., de Araujo, E.S., Stabellini, R., Fraga, A.M., de Souza, J.E., Sumita, D.R., Camargo, A.A., and Pereira, L.V. (2010) Random X inactivation and extensive mosaicism in human placenta revealed by analysis of allele-specific gene expression along the X chromosome. *PLoS One* 5(6), e10947

Nagano, M., Watson, D.J., Ryu, B.Y., Wolfe, J.H., and Brinster, R.L. (2002) Lentiviral vector transduction of male germ line stem cells in mice. *FEBS Lett* 524(1-3), 111-5

Namekawa, S.H., Payer, B., Huynh, K.D., Jaenisch, R., and Lee, J.T. (2010) Two-step imprinted X inactivation: repeat versus genic silencing in the mouse. *Mol Cell Biol* 30(13), 3187-205

Namekawa, S.H., VandeBerg, J.L., McCarrey, J.R., and Lee, J.T. (2007) Sex chromosome silencing in the marsupial male germ line. *Proc Natl Acad Sci U S A* 104(23), 9730-5

Navarro, P., Chambers, I., Karwacki-Neisius, V., Chureau, C., Morey, C., Rougeulle, C., and Avner, P. (2008) Molecular coupling of Xist regulation and pluripotency. *Science* 321(5896), 1693-5

Navarro, P., Moffat, M., Mullin, N.P., and Chambers, I. (2011) The X-inactivation trans-activator Rnf12 is negatively regulated by pluripotency factors in embryonic stem cells. *Hum Genet* 130(2), 255-64

Navarro, P., Oldfield, A., Legoupi, J., Festuccia, N., Dubois, A., Attia, M., Schoorlemmer, J., Rougeulle, C., Chambers, I., and Avner, P. (2010) Molecular coupling of Tsix regulation and pluripotency. *Nature* 468(7322), 457-60

Navarro, P., Page, D.R., Avner, P., and Rougeulle, C. (2006) Tsix-mediated epigenetic switch of a CTCF-flanked region of the Xist promoter determines the Xist transcription program. *Genes Dev* 20(20), 2787-92

Nichols, J., and Smith, A. (2009) Naive and primed pluripotent states. *Cell Stem Cell* 4(6), 487-92

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998) Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* 95(3), 379-91

Niwa, H., Miyazaki, J., and Smith, A.G. (2000) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* 24(4), 372-6

Nolen, L.D., Gao, S., Han, Z., Mann, M.R., Gie Chung, Y., Otte, A.P., Bartolomei, M.S., and Latham, K.E. (2005) X chromosome reactivation and regulation in cloned embryos. *Dev Biol* 279(2), 525-40

Norris, D.P., Brockdorff, N., and Rastan, S. (1991) Methylation status of CpG-rich

islands on active and inactive mouse X chromosomes. *Mamm Genome* 1(2), 78-83

Norris, D.P., Patel, D., Kay, G.F., Penny, G.D., Brockdorff, N., Sheardown, S.A., and Rastan, S. (1994) Evidence that random and imprinted Xist expression is controlled by preemptive methylation. *Cell* 77(1), 41-51

Nowak-Imialek, M., and Niemann, H. (2012) Pluripotent cells in farm animals: state of the art and future perspectives. *Reprod Fertil Dev* 25(1), 103-28

Ogawa, Y., and Lee, J.T. (2003) Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. *Mol Cell* 11(3), 731-43

Ohhata, T., Senner, C.E., Hemberger, M., and Wutz, A. (2011) Lineage-specific function of the noncoding Tsix RNA for Xist repression and Xi reactivation in mice. *Genes Dev* 25(16), 1702-15

Ohhata, T., and Wutz, A. (2013) Reactivation of the inactive X chromosome in development and reprogramming. *Cell Mol Life Sci* 70(14), 2443-61

Okamoto, I., Arnaud, D., Le Baccon, P., Otte, A.P., Disteche, C.M., Avner, P., and Heard, E. (2005) Evidence for de novo imprinted X-chromosome inactivation independent of meiotic inactivation in mice. *Nature* 438(7066), 369-73

Okamoto, I., and Heard, E. (2009) Lessons from comparative analysis of X-chromosome inactivation in mammals. *Chromosome Res* 17(5), 659-69

Okamoto, I., Otte, A.P., Allis, C.D., Reinberg, D., and Heard, E. (2004) Epigenetic

dynamics of imprinted X inactivation during early mouse development. *Science* 303(5658), 644-9

Okamoto, I., Patrat, C., Thepot, D., Peynot, N., Fauque, P., Daniel, N., Diabangouaya, P., Wolf, J.P., Renard, J.P., Duranthon, V., and Heard, E. (2011) Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. *Nature* 472(7343), 370-4

Okamoto, K., Okazawa, H., Okuda, A., Sakai, M., Muramatsu, M., and Hamada, H. (1990) A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. *Cell* 60(3), 461-72

Pang, K.C., Frith, M.C., and Mattick, J.S. (2006) Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *Trends Genet* 22(1), 1-5

Panning, B., and Jaenisch, R. (1996) DNA hypomethylation can activate Xist expression and silence X-linked genes. *Genes Dev* 10(16), 1991-2002

Papamichos, S.I., Kotoula, V., Tarlatzis, B.C., Agorastos, T., Papazisis, K., and Lambropoulos, A.F. (2009) OCT4B1 isoform: the novel OCT4 alternative spliced variant as a putative marker of stemness. *Mol Hum Reprod* 15(5), 269-70

Park, C.H., Jeong, Y.H., Jeong, Y.I., Lee, S.Y., Jeong, Y.W., Shin, T., Kim, N.H., Jeung, E.B., Hyun, S.H., Lee, C.K., Lee, E., and Hwang, W.S. (2012) X-linked gene transcription patterns in female and male in vivo, in vitro and cloned porcine individual blastocysts. *PLoS One* 7(12), e51398

Park, C.H., Kim, H.S., Lee, S.G., and Lee, C.K. (2009) Methylation status of

differentially methylated regions at Igf2/H19 locus in porcine gametes and preimplantation embryos. *Genomics* 93(2), 179-86

Park, C.H., Uh, K.J., Mulligan, B.P., Jeung, E.B., Hyun, S.H., Shin, T., Ka, H., and Lee, C.K. (2011) Analysis of imprinted gene expression in normal fertilized and uniparental preimplantation porcine embryos. *PLoS One* 6(7), e22216

Park, J.-K., Kim, H.-S., Uh, K.-J., Choi, K.-H., Kim, H.-M., Lee, T., Yang, B.-C., Kim, H.-J., Ka, H.-H., Kim, H., and Lee, C.-K. (2013a) Primed Pluripotent Cell Lines Derived from Various Embryonic Origins and Somatic Cells in Pig. *PLoS One* 8(1), e52481

Park, J.K., Kim, H.S., Uh, K.J., Choi, K.H., Kim, H.M., Lee, T., Yang, B.C., Kim, H.J., Ka, H.H., Kim, H., and Lee, C.K. (2013b) Primed pluripotent cell lines derived from various embryonic origins and somatic cells in pig. *PLoS One* 8(1), e52481

Park, Y., and Kuroda, M.I. (2001) Epigenetic aspects of X-chromosome dosage compensation. *Science* 293(5532), 1083-5

Patrat, C., Okamoto, I., Diabangouaya, P., Vialon, V., Le Baccon, P., Chow, J., and Heard, E. (2009) Dynamic changes in paternal X-chromosome activity during imprinted X-chromosome inactivation in mice. *Proc Natl Acad Sci U S A* 106(13), 5198-203

Payer, B., Lee, J.T., and Namekawa, S.H. (2011) X-inactivation and X-reactivation: epigenetic hallmarks of mammalian reproduction and pluripotent stem cells. *Hum Genet* 130(2), 265-80

Payer, B., Rosenberg, M., Yamaji, M., Yabuta, Y., Koyanagi-Aoi, M., Hayashi, K., Yamanaka, S., Saitou, M., and Lee, J.T. (2013) Tsix RNA and the germline factor,

PRDM14, link X reactivation and stem cell reprogramming. *Mol Cell* 52(6), 805-18

Pehrson, J.R., and Fried, V.A. (1992) MacroH2A, a core histone containing a large nonhistone region. *Science* 257(5075), 1398-400

Peippo, J., Farazmand, A., Kurkilahti, M., Markkula, M., Basrur, P.K., and King, W.A. (2002) Sex-chromosome linked gene expression in in-vitro produced bovine embryos. *Mol Hum Reprod* 8(10), 923-9

Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., and Brockdorff, N. (1996) Requirement for Xist in X chromosome inactivation. *Nature* 379(6561), 131-7

Pesce, M., Wang, X., Wolgemuth, D.J., and Scholer, H. (1998) Differential expression of the Oct-4 transcription factor during mouse germ cell differentiation. *Mech Dev* 71(1-2), 89-98

Pillet, N., Bonny, C., and Schorderet, D.F. (1995) Characterization of the promoter region of the mouse Xist gene. *Proc Natl Acad Sci U S A* 92(26), 12515-9

Plath, K., Fang, J., Mlynarczyk-Evans, S.K., Cao, R., Worringer, K.A., Wang, H., de la Cruz, C.C., Otte, A.P., Panning, B., and Zhang, Y. (2003) Role of histone H3 lysine 27 methylation in X inactivation. *Science* 300(5616), 131-5

Pugacheva, E.M., Tiwari, V.K., Abdullaev, Z., Vostrov, A.A., Flanagan, P.T., Quitschke, W.W., Loukinov, D.I., Ohlsson, R., and Lobanenko, V.V. (2005) Familial cases of point mutations in the XIST promoter reveal a correlation between CTCF binding and pre-emptive choices of X chromosome inactivation. *Hum Mol Genet* 14(7), 953-65

Pullirsch, D., Hartel, R., Kishimoto, H., Leeb, M., Steiner, G., and Wutz, A. (2010) The Trithorax group protein Ash2l and Saf-A are recruited to the inactive X chromosome at the onset of stable X inactivation. *Development* 137(6), 935-43

Rastan, S. (1983) Non-random X-chromosome inactivation in mouse X-autosome translocation embryos--location of the inactivation centre. *J Embryol Exp Morphol* 78, 1-22

Rastan, S., and Robertson, E.J. (1985) X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. *J Embryol Exp Morphol* 90, 379-88

Richardson, B.J., Czuppon, A.B., and Sharman, G.B. (1971) Inheritance of glucose-6-phosphate dehydrogenase variation in kangaroos. *Nat New Biol* 230(13), 154-5

Romito, A., and Rougeulle, C. (2011) Origin and evolution of the long non-coding genes in the X-inactivation center. *Biochimie* 93(11), 1935-42

Rosner, M.H., Vigano, M.A., Ozato, K., Timmons, P.M., Poirier, F., Rigby, P.W., and Staudt, L.M. (1990) A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature* 345(6277), 686-92

Rossant, J. (2008) Stem cells and early lineage development. *Cell* 132(4), 527-31

Sado, T., Fenner, M.H., Tan, S.S., Tam, P., Shioda, T., and Li, E. (2000) X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation. *Dev Biol* 225(2), 294-303

Sado, T., Hoki, Y., and Sasaki, H. (2005) Tsix silences Xist through modification of chromatin structure. *Dev Cell* 9(1), 159-65

Sado, T., and Sakaguchi, T. (2013) Species-specific differences in X chromosome inactivation in mammals. *Reproduction* 146(4), R131-9

Sado, T., Wang, Z., Sasaki, H., and Li, E. (2001) Regulation of imprinted X-chromosome inactivation in mice by Tsix. *Development* 128(8), 1275-86

Sakurai, N., Fujii, T., Hashizume, T., and Sawai, K. (2013) Effects of downregulating oct-4 transcript by RNA interference on early development of porcine embryos. *J Reprod Dev* 59(4), 353-60

Sarma, K., Levasseur, P., Aristarkhov, A., and Lee, J.T. (2010) Locked nucleic acids (LNAs) reveal sequence requirements and kinetics of Xist RNA localization to the X chromosome. *Proc Natl Acad Sci U S A* 107(51), 22196-201

Scholer, H.R., Ruppert, S., Suzuki, N., Chowdhury, K., and Gruss, P. (1990) New type of POU domain in germ line-specific protein Oct-4. *Nature* 344(6265), 435-9

Schulz, E.G., and Heard, E. (2013) Role and control of X chromosome dosage in mammalian development. *Curr Opin Genet Dev* 23(2), 109-15

Sharman, G.B. (1971) Late DNA replication in the paternally derived X chromosome of female kangaroos. *Nature* 230(5291), 231-2

Sheardown, S.A., Duthie, S.M., Johnston, C.M., Newall, A.E., Formstone, E.J., Arkell, R.M., Nesterova, T.B., Alghisi, G.C., Rastan, S., and Brockdorff, N. (1997a) Stabilization of Xist RNA mediates initiation of X chromosome inactivation. *Cell* 91(1), 99-107

Sheardown, S.A., Newall, A.E., Norris, D.P., Rastan, S., and Brockdorff, N. (1997b) Regulatory elements in the minimal promoter region of the mouse Xist gene. *Gene* 203(2), 159-68

Shin, J., Bossenz, M., Chung, Y., Ma, H., Byron, M., Taniguchi-Ishigaki, N., Zhu, X., Jiao, B., Hall, L.L., Green, M.R., Jones, S.N., Hermans-Borgmeyer, I., Lawrence, J.B., and Bach, I. (2010) Maternal Rnf12/RLIM is required for imprinted X-chromosome inactivation in mice. *Nature* 467(7318), 977-81

Shin, J., Wallingford, M.C., Gallant, J., Marcho, C., Jiao, B., Byron, M., Bossenz, M., Lawrence, J.B., Jones, S.N., Mager, J., and Bach, I. (2014) RLIM is dispensable for X-chromosome inactivation in the mouse embryonic epiblast. *Nature* 511(7507), 86-9

Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T.B., Webster, Z., Peters, A.H., Jenuwein, T., Otte, A.P., and Brockdorff, N. (2003) Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Dev Cell* 4(4), 481-95

Silva, S.S., Rowntree, R.K., Mekhoubad, S., and Lee, J.T. (2008) X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. *Proc Natl Acad Sci U S A* 105(12), 4820-5

Singer-Sam, J., Grant, M., LeBon, J.M., Okuyama, K., Chapman, V., Monk, M., and Riggs, A.D. (1990) Use of a HpaII-polymerase chain reaction assay to study DNA

methylation in the P_{gk}-1 CpG island of mouse embryos at the time of X-chromosome inactivation. *Mol Cell Biol* 10(9), 4987-9

Soma, M., Fujihara, Y., Okabe, M., Ishino, F., and Kobayashi, S. (2014) Ftx is dispensable for imprinted X-chromosome inactivation in preimplantation mouse embryos. *Sci Rep* 4, 5181

Su, J.M., Yang, B., Wang, Y.S., Li, Y.Y., Xiong, X.R., Wang, L.J., Guo, Z.K., and Zhang, Y. (2011) Expression and methylation status of imprinted genes in placentas of deceased and live cloned transgenic calves. *Theriogenology* 75(7), 1346-59

Sun, B.K., Deaton, A.M., and Lee, J.T. (2006) A transient heterochromatic state in Xist preempts X inactivation choice without RNA stabilization. *Mol Cell* 21(5), 617-28

Sun, S., Del Rosario, B.C., Szanto, A., Ogawa, Y., Jeon, Y., and Lee, J.T. (2013) Jpx RNA activates Xist by evicting CTCF. *Cell* 153(7), 1537-51

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131(5), 861-72

Takahashi, K., and Yamanaka, S. (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126(4), 663-76

Takeda, J., Seino, S., and Bell, G.I. (1992) Human Oct3 gene family: cDNA sequences, alternative splicing, gene organization, chromosomal location, and expression at low levels in adult tissues. *Nucleic Acids Res* 20(17), 4613-20

Tanasijevic, B., Dai, B., Ezashi, T., Livingston, K., Roberts, R.M., and Rasmussen, T.P. (2009) Progressive accumulation of epigenetic heterogeneity during human ES cell culture. *Epigenetics* 4(5), 330-8

Tapia, N., Reinhardt, P., Duemmler, A., Wu, G., Arauzo-Bravo, M.J., Esch, D., Greber, B., Cojocaru, V., Rascon, C.A., Tazaki, A., Kump, K., Voss, R., Tanaka, E.M., and Scholer, H.R. (2012) Reprogramming to pluripotency is an ancient trait of vertebrate Oct4 and Pou2 proteins. *Nat Commun* 3, 1279

Tchieu, J., Kuoy, E., Chin, M.H., Trinh, H., Patterson, M., Sherman, S.P., Aimiwu, O., Lindgren, A., Hakimian, S., Zack, J.A., Clark, A.T., Pyle, A.D., Lowry, W.E., and Plath, K. (2010) Female human iPSCs retain an inactive X chromosome. *Cell Stem Cell* 7(3), 329-42

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D. (2007) New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448(7150), 196-9

Tian, D., Sun, S., and Lee, J.T. (2010) The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. *Cell* 143(3), 390-403

Tribioli, C., Tamanini, F., Patrosso, C., Milanesi, L., Villa, A., Pergolizzi, R., Maestrini, E., Rivella, S., Bione, S., Mancini, M., and et al. (1992) Methylation and sequence analysis around EagI sites: identification of 28 new CpG islands in XQ24-XQ28. *Nucleic Acids Res* 20(4), 727-33

Turner, J.M., Mahadevaiah, S.K., Ellis, P.J., Mitchell, M.J., and Burgoyne, P.S. (2006) Pachytene asynapsis drives meiotic sex chromosome inactivation and leads to substantial

postmeiotic repression in spermatids. *Dev Cell* 10(4), 521-9

van den Berg, I.M., Laven, J.S., Stevens, M., Jonkers, I., Galjaard, R.J., Gribnau, J., and van Doorninck, J.H. (2009) X chromosome inactivation is initiated in human preimplantation embryos. *Am J Hum Genet* 84(6), 771-9

Van Soom, A., Wrathall, A.E., Herrler, A., and Nauwynck, H.J. (2010) Is the zona pellucida an efficient barrier to viral infection? *Reprod Fertil Dev* 22(1), 21-31

Vigneau, S., Augui, S., Navarro, P., Avner, P., and Clerc, P. (2006) An essential role for the DXPas34 tandem repeat and Tsix transcription in the counting process of X chromosome inactivation. *Proc Natl Acad Sci U S A* 103(19), 7390-5

Wang, J., Mager, J., Chen, Y., Schneider, E., Cross, J.C., Nagy, A., and Magnuson, T. (2001) Imprinted X inactivation maintained by a mouse Polycomb group gene. *Nat Genet* 28(4), 371-5

Wang, X., and Dai, J. (2010) Concise review: isoforms of OCT4 contribute to the confusing diversity in stem cell biology. *Stem Cells* 28(5), 885-93

Wang, X., Zhao, Y., Xiao, Z., Chen, B., Wei, Z., Wang, B., Zhang, J., Han, J., Gao, Y., Li, L., Zhao, H., Zhao, W., Lin, H., and Dai, J. (2009) Alternative translation of OCT4 by an internal ribosome entry site and its novel function in stress response. *Stem Cells* 27(6), 1265-75

Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Paabo, S., Rebhan, M., and Schubeler, D. (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* 39(4), 457-66

Wernersson, R., and Pedersen, A.G. (2003) RevTrans: Multiple alignment of coding DNA from aligned amino acid sequences. *Nucleic Acids Res* 31(13), 3537-9

Wilmut, I., Beaujean, N., de Sousa, P.A., Dinnyes, A., King, T.J., Paterson, L.A., Wells, D.N., and Young, L.E. (2002) Somatic cell nuclear transfer. *Nature* 419(6907), 583-6

Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J., and Campbell, K.H. (1997) Viable offspring derived from fetal and adult mammalian cells. *Nature* 385(6619), 810-3

Wrenzycki, C., Lucas-Hahn, A., Herrmann, D., Lemme, E., Korsawe, K., and Niemann, H. (2002) In vitro production and nuclear transfer affect dosage compensation of the X-linked gene transcripts G6PD, PGK, and Xist in preimplantation bovine embryos. *Biol Reprod* 66(1), 127-34

Wutz, A. (2011) Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nat Rev Genet* 12(8), 542-53

Wutz, A., and Jaenisch, R. (2000) A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol Cell* 5(4), 695-705

Wutz, A., Rasmussen, T.P., and Jaenisch, R. (2002) Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat Genet* 30(2), 167-74

Xue, F., Tian, X.C., Du, F., Kubota, C., Taneja, M., Dinnyes, A., Dai, Y., Levine, H., Pereira, L.V., and Yang, X. (2002) Aberrant patterns of X chromosome inactivation in bovine clones. *Nat Genet* 31(2), 216-20

Yamaji, M., Seki, Y., Kurimoto, K., Yabuta, Y., Yuasa, M., Shigeta, M., Yamanaka, K., Ohinata, Y., and Saitou, M. (2008) Critical function of Prdm14 for the establishment of the germ cell lineage in mice. *Nat Genet* 40(8), 1016-22

Yamaji, M., Ueda, J., Hayashi, K., Ohta, H., Yabuta, Y., Kurimoto, K., Nakato, R., Yamada, Y., Shirahige, K., and Saitou, M. (2013) PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. *Cell Stem Cell* 12(3), 368-82

Yen, Z.C., Meyer, I.M., Karalic, S., and Brown, C.J. (2007) A cross-species comparison of X-chromosome inactivation in Eutheria. *Genomics* 90(4), 453-63

Yoshioka, K., Suzuki, C., Tanaka, A., Anas, I.M., and Iwamura, S. (2002) Birth of piglets derived from porcine zygotes cultured in a chemically defined medium. *Biol Reprod* 66(1), 112-9

Young, L.E., Sinclair, K.D., and Wilmut, I. (1998) Large offspring syndrome in cattle and sheep. *Rev Reprod* 3(3), 155-63

Zhao, J., Sun, B.K., Erwin, J.A., Song, J.J., and Lee, J.T. (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* 322(5902), 750-6

Zhou, W., Xiang, T., Walker, S., Farrar, V., Hwang, E., Findeisen, B., Sadeghieh, S., Arenivas, F., Abruzzese, R.V., and Polejaeva, I. (2008) Global gene expression analysis of bovine blastocysts produced by multiple methods. *Mol Reprod Dev* 75(5), 744-58

Abstract in Korean

X 염색체 불활성화 현상은 암컷 태반동물의 배아 발달 과정에서 특이적으로 확인되는 후생유전학적 특성으로, 이는 수컷과 암컷 성체에서의 X 염색체 내 유전자 발현 양이 같아지도록 하기 위해 발생한다. 생쥐 모델에서의 연구는 이 현상의 분자생물학적 메커니즘 및 진행 과정을 밝혀왔다. 또한 최근의 연구는 이 현상이 배아 발달 과정 및 줄기세포에서의 만능성과 연관성이 있음을 확인함으로써 해당 연구의 중요성을 강조해왔다. 하지만 이러한 중요성에도 불구하고, 지난 50년간의 해당 분야의 연구는 생쥐 모델에만 국한되어 있는 상황이다. 최근 X 염색체 불활성화 기작과 dl 현상의 조절인자가 중간에 다양함이 밝혀짐에 따라 해당 현상의 연구는 좀 특이적으로 진행이 되어야 된다는 의견이 제시되고 있다. 이에 본 연구에서는 돼지에서의 X 염색체 불활성화 현상에 대하여 알아보고자 1) X 염색체 불활성화 현상의 주요 인자인 XIST 유전자와 X 염색체 불활성화 센터의 확립, 2) 배아 발달과정에서의 X 염색체 불활성화 현상의 확인, 3) 해당 현상과 만능성 유전자와의 연관성을 알아보고자 하였다.

첫 번째 연구는 돼지의 XIST 유전자 확립을 목표로 하였다. 이 유전자는 X 염색체 불활성화 현상의 주된 유발 인자로 X 염색체를 코팅함으로써 유전자의 발현을 억제한다. 비록 인간과 생쥐의 XIST 유전자는 20여년 전에 확립이 되었지만, 최근까지 돼지에서는 이 유전자가 확립되지 않았으며, 이를 확인하기 위해 중간 XIST 유전자 염기서열의 비교 분석 및 성별에 따른 배아 섬유아세포에서의 발현 양상을 확인하였다. 그 결과 25Kb의 전사체를 생산해내는 유전자를 확립할 수 있으며 이 유전자의 엑손 구성은 인간과 생쥐 XIST 유전자의 엑손 구성과 유사하였으나 염기서열의 유사성은 낮음을 확인할 수 있었다. 이 유전자는 암컷의 배아 섬유아세포에서 특이적으로 발현하는 것을 확인할 수 있었으며, 유전자 내의 4개의 반복서열 중 첫 번째와 두 번째 반복 서열은 생쥐의 Repeat A 및 Repeat B 구간과 유사성이 높음을 알 수 있었다. 또한 이 유전자의 프로모터 위치에 존재하는 CpG 염기서열은 활성화된 X 염색체 만을 보유하고 있는 수컷의 배아 섬유아세포에서 높은 수준으로 메틸화 되어있는 것을 알 수 있었다. 이를 통해 첫 번째 연구에서는 확립된 유전자가 돼지의 XIST 유전자임을 확인할 수 있었다.

XIST 유전자와 가까이 위치하는 유전자들은 최근 X 염색체 불활성화 현상에 관여한다는 연구결과가 생쥐모델에서 확인되었다. 이 유전자들을

포함하고 있는 게놈상의 특정 위치를 X 염색체 불활성화 센터라고 한다. 하지만 이 센터 내에 존재하는 다양한 유전자들이 XIST 유전자를 활성화 시키거나 발현을 억제함으로써 불활성화 현상에 개입된다는 점에도 불구하고 돼지에서 이 센터는 현재까지 확인되지 않고 있다. 따라서 두 번째 연구에서는 돼지에서의 X 염색체 불활성화 센터를 확립하고, 이 센터 내 유전자의 후생유전학적 변화를 배아 발달 단계에서 확인함으로써 돼지에서의 X 염색체 불활성화 현상의 진행을 확인해 보고자 하였다. X 염색체 불활성화 센터는 인간과 생쥐에서 각각 1.1 Mb와 0.5Mb의 게놈 구간상에 존재하고 단백질을 암호화 하거나 암호화 하지 않는 유전자들로 구성되어 있다. 단백질을 암호화 하지 않는 유전자의 유대류와 진수류의 진화 분기점에서 원시 단백질 유전자의 위유전자 화에 의해 생성된 것으로 알려져 있다. X 염색체 불활성화 센터를 구성하는 유전자의 배열은 진수류 중 간에 보존되어 있으나 각 유전자의 염기서열, 특히 단백질 미 암호화 유전자의 경우 염기서열의 유사성이 상당히 낮다. 이러한 특성을 토대로 인간의 X 염색체 불활성화 센터의 유전자 염기서열과 돼지의 게놈을 비교해본 결과, 약 1.1 Mb의 게놈 구간이 인간의 X 염색체 불활성화 센터와 상동성을 가짐을 확인할 수 있었고, 이 구간을 돼지의 X 염색체 불활성화 센터라고 결정하였다. 이 구간은 XIST 유전자를 포함하고 있었으며, 단백질 미 암호화 유전자의 경우 상대적으로 단백질

암호화 유전자에 비하여 각각의 인간에서의 대응 유전자와의 염기서열 유사성이 낮음을 알 수 있었다. 이는 확립된 게놈 구간이 진수류의 X 염색체 불활성화 센터가 보유하고 있는 진화적 특성을 공유한다는 것을 의미한다. 이 센터 내에 존재하는 유전자 중, CHIC1, XIST, LOC102165544, RLIM 유전자가 돼지의 암컷 배아 발달 과정 중 안정적으로 발현하는 것을 알 수 있었다. 이 중 CHIC1과 RLIM 유전자는 상실배에서 배반포로 발달이 진행함에 따라 유전자 발현이 감소하였으며 배반포에서 성별에 따른 유전자의 발현양에 차이가 나지 않았다. 이는 발달 과정 중 유전자의 발현 감소가 암컷과 수컷 배반포에서의 유전자 발현 양 균등화를 유발하였음을 의미한다. 하지만 단백질 미 암호화 유전자인 XIST와 LOC102165544는 상실배 이후 발현양이 증가함을 보였고 배반포에서 성별에 따른 발현양의 차이가 존재하는 것이 확인되었다. 암컷 배반포에서 CHIC1 유전자의 프로모터는 하나의 대립 유전자가 메틸화가 되어있음을 확인 하였으나 XIST를 포함한 나머지 유전자의 경우 프로모터에서의 메틸화가 확인되지 않았다. 이는 돼지의 배반포에서는 일부 유전자의 유전자량 보정이 확립 되었지만 염색체 수준에서의 불활성화는 확립되지 않았음을 의미한다.

돼지에서의 X 염색체 불활성화 과정을 확인한 뒤, 만능성 유지에 필수적이며 최근 XIST 유전자의 발현을 조절한다고 알려진 OCT4 유전자가 돼지의 배아 발달 과정 중 X 염색체 불활성화 센터에 존재하는 유전자의 발현에 어떠한 영향을 미치는지 알아보고자 하였다. 이에 앞서, 최근 OCT4 이형체의 발현이 정확한 분석에 부정적인 영향을 미친다는 기존의 보고에 따라 돼지의 OCT4 이형체 존재 여부와 발현에 대해 확인을 세 번째 연구에서 진행하고자 하였다. 돼지의 OCT4 유전자는 기존에 확립된 전사체 (OCT4A) 이외에 두 개의 이형체 (OCT4B, OCT4B1)를 추가적으로 전사함이 확인되었고, 이 이형체들이 OCT4A를 구성하는 엑손을 공유하고 있음이 확인되었다. 또한 OCT4B 전사체는 단백질을 생산하며 성체 조직에서 발현되는 것을 알 수 있었고, 배아의 세포질에서 특이적으로 발현되는 것을 확인할 수 있었다.

이후 네 번째 연구에서는 OCT4A와 X 염색체 불활성화 센터 내 유전자의 발현의 연관성을 인간의 OCT4A 유전자를 돼지의 배아에서 과발현 시킴으로써 확인해 보고자 하였다. 기존의 생쥐 모델에서의 연구 결과와는 다르게, OCT4A의 과발현은 돼지의 배반포에서 XIST 유전자의 발현 증가를 유발하였다. 마지막 연구에서는 배반포에서 만능성 유전자와 X 염색체 불활성화 센터 내 유전자 발현량의 상관관계를 분석하였다. 분석

결과 만능성 유전자인 NANOG와 REX1이 XIST와 RLIM 유전자와 각각 연관성이 있을 수 있음을 알 수 있었다. 비록 추가적인 실험이 필요하지만, 일련의 결과를 통해 돼지의 배아 발달과정에서 만능성 인자들이 X 염색체 불활성화 현상과 관련이 있음을 확인할 수 있었다.

이상의 연구에서는 돼지에서 X 염색체 불활성화 현상의 핵심 인자를 확립하고 이를 기반으로 배아 발달 과정에서의 그 현상을 확인하였다. 또한 다양한 인자들이 돼지의 배아 발달 과정 중 X 염색체 불활성화 현상에 관여한다는 사실을 알 수 있었다. 본 연구는 돼지의 X 염색체 불활성화 현상에 대한 새로운 이해를 제시하였고 돼지의 발생학과 줄기세포 연구에 기여할 것이다.

주요어 : X 염색체 불활성화, X 염색체 불활성화 센터, XIST, OCT4, 착상 전 배아, 배반포, 유전자 발현, 돼지

학번 : 2009 - 21275